



FIGURE 1

FIGURE 2, Sheet 3 of 40

AAGACGACGACTGCGAAGAAGAAGGCGACTTTCTTCCGAGTTAAATCCAACGCACTCTCAACTACTTCTATTCCAACACGGCAACAGATAGCTGCAC
 AGAAGACGACGACGTTTGTAGTCTGTAGAAGAATAACAACAGAGTAAGCGGTTCTTCCACCACAGCCGGAGACAGAGTTCTTGAAGAGGATCTTCTC
 TCTACTGTATCTCCGAACGAAAAGAGGAACCTCTGCCGCCCTCGCCGCACTCACCATATCCCGGCACTCTCTTTCAACGCTCTATCTGCAAAAACAAAGT
 TGGGAGAAAATGGACGTTTCTTCCATAAGAGCACTATTGACTACCACAACTTGAAGATATGGACGATCTCCAGCGGCCACCTACAAGGATCGTATGG
 AGACGGAATTGGTCTCGAGATGGCTAAGAAGGAGGGAAGGTACGTCCGATCGTTGGCCACCATGGACGAATTGGAGGTACCTGAAGAACCAGCCACTTG
 CTACACTTGGCGCTACACCTTTATTAGACGCGAGGCAACCCCAACAAACGCAAGTCAATATTCAGAGAGCCTTGGCGTTACCCAGAAGCTTCTCCCGAT
 GCACCATCCCCCGTCCGTTTAGAAGAGCTTGTGACGTGCCAGAGGAGCGAGTTTTTTCACCTACCTCCCTACGACGACGGATCTTCTACATCGTCTT
 CACAAGCCGAATGTGAAGATGATTATCTCCACCATACGACCCATCAGAAAATCCACAGAGGTCCCAAGTGTGTGATTATTGTACCACAGCTCAAGTCTT
 CAGTCTATGACGGATCAGCCAGGCGCAACCTCATAAAAATCTGAAGAGGAGAAGAAGGCCCTGGGCTTGGCGCTCGCAACAACTTTAGCTACTAG
 GGTGTACAAGAAGAAAATAGGATTGATTGTATCGATGACGAACCTTGTGATGTTTTTGACCATCTTTTGGCTTGGCTGAGTGTACATAGCTTTATATTTCG
 CAATAAAACAGACCCCGAATGTATTTTGTGTATTCTTTATTACTGCCATTTTTATGGTTAAACGGAAAAAATTTTGAAGTTTGTGAGATGGAGAT
 GAAGGGAAGAAGGCGCTAGTTTATATAGTTGGTATATTATTTCTGGTATACATTTCTAGCTCCTCTGTGAATCTGACGTTGGGTCGACCCAGCG
 GTCCACCTTCCGAATGTACATAAGGCCCTAGTCAGCGGTCCACCCCTAACTGGAGTGAGCTGAAAAAATTTTGAAGTTTGTGAGATGGAGATGAA
 GGGTAAAAAGAGGGTTGAAGGATAGCTGAGGCCACCTGGGTGACCGGTGTTCCAGAAACGCTGTCTGTCTCCAGAAATGTTGTAGAGATTCTGGACA
 AGTCATTTCTGGAAGGGTGAATTTATTAAGTTGGTATATTATTTCTGGTATATATTCTAGCCCTCTCTGTAACTGACGTTGGGTCGACCCAGC
 GGTCCACCTTCCGAATGTACATAAGGCCCTAGTCAGCGGTCCACCCCTAACTGGAGTGAGCTGAAAAAATTTTGAAGTTTGTGAGATGGAGATGAA
 GGGTAAAAAGGCTAGTAAAGAGGTTGTGTAACCTCGGATACCAACGAGAAATATATCAAAAGGGATGGAAACAACTATTCTTTTATACAGATAGTTTAA
 TTGTACCAAAACTCAGAAATCTCATACATGACAAATCTGAATGCCATTCGTTGGCCCTCAATAACGTACGACACTGGACCTGGGAACCTCCACCTA
 CATGTTTACATACAGAAATTTTGGCTTAGATGCATTCAACAAACACTTTTACCACATTGGCATTCTCTCTCAGTTTGTICACACATCCCAAGAA
 AGGCCATATTCTCCATAGAAATCAGATCCATCAACAAATATATAGAAATGTACCAATTTCTGCGATCCAGTCCAGTCCAGTCTTGTATCTTGTATCTGACAAA
 AACCATGTTTCCGAAGATTCTTGTGTCCTCCATGTATGGGAAAGATACAGTATGGTCCACAGTCTTGTATCTCTCCCAACAGCTCGTTCGAAAAA
 TCCCTAATAAAGAAAGCAGGCAATTCGACTGTATCGAAACCAATGTTGAGTTTCTCTGATTCTTCTCATTGTTGGGAGGAAGAAATCGCCAAAT
 GTGTATCTGCTTCAATGCTCCACACTTCCCTTTATGGATCAGCGGTTCTCCATAATCTCCTTGGCAGAAATGGGGAGTAGGCCCTCGAAGAACAGGAT
 ATTTGGGCGGGTAGGAACCTTGGCAGTCAAGTCCATGTTGGCAAGAGATCAGCAGGGTCTTCAAGACTGAGGCCATTTTGGTGTGCTGCTGAG
 AGAGGAGCAAGTAGTAGTGTGGTTACTAGCAGGAGCTCACAGTCTTATATATAGACTAACACCCCGCCCGCCCAAACTCTCCACATCTTATACCG
 GGTGCGGTAGAACCCAGAACCGTTGACTCAGATCTGGTACATTTCTGTTATACATTTCTAGCCCTCTGCTGAATCTGACATTGGGTCGACCCA
 CGCGTCCACCTTCCGAATGTGACATTCGATCGACTCAGCGGTCCATCCCTAACTGGAGTGACCCAGAAAAATTTTGAAGTTTGTGAGATGAAGAT
 GAAGGGGAAAAGGTTAGTGAATACAGGCACTTCTGCTTCTCCAGAAATGATTCTTAAAGATTCTGTGCAATTTCTGGTATACATTTCTAGCCCC
 TCTGTTAATCTGACATTGGGCGGACCCAGCGGTCCACCTCGGAACCTGACATTCGATCGACTCAGCGGTCCACCCCTAACTGGAGTGACCCAGAAA
 AATTTTGAAGTTTGTGAGATGAAGTGAAGGGGAAAAGAGGGTCTAGGGTGGCTCCAGACATCTCTAGACTGGCTAGAGCCCGAAGAACATCC
 AACTAGTTTCTGGGACATTTTCTCCACGTAGCGCAAGAAATAGTAACTAACAGATATTTGCAACATTTTGCACACATTTCTACCATCTCGGCAATA
 CACTGTTGTGTTGCTCCGATAGCTTAACATGGCCCAAACTCTTCCAGAAATTTGCTCCAGTTATCAAGACTGAGAAGAGGAAAGAAAGGAGGAT
 AACATGACGACCTTTACGGCAGATTGATTTAGAGATAGAAGACATTAATCTGCTCAGTCAAGAACTGTGTTTCGAGGAGAAAGAACTGGATCTGC
 ACATGATCGAGTATACAAAGTACTACGTACGGGAACCCATACAGTACCGTCCGCCCCAATAGAACACATCGAGGATTGGCCCTCTCAATGGATCAAGGT
 GAAGTAGGAACATGCTTCCCTCTGCGACCTTGGAGAGACTGAAGAAACCCCATCGACAGTGGGAGTGGCGTTCTGATCTCAACTACAACTGAAG
 AAGATGGATGACCCACCTTTACAACGACGAAGATATATAAGAACTGCAACAAATTTGAAGGAGAAACAAAGTGGGTAAAGAAAGCTCGGCAAGA
 ATACTTCAGACAAGCTCTAGAGACATTGATGATGCCATTTCTATAAACAATATTCCAATTTTATTTTTCAGGAGGATATGGAGGAAGGATTGTG
 CACAACTCCACACATTTATTAATATGGTACACCTAAAAAGGTGCTGTTTGTGATATAAAGATTGATGTTTATAAATGAACGATTTATTTATTA
 TACCAAAAATTTTATGATGATTTTCAATGGGATGGCTTCTGCTACCAATCTGGCCATGTGTGCTGTTCTATGCTAGCCCTCTTACCCCTTCATC
 TCCATCTCAAAAATTTTGAAGAAATTTTCTGGGTGCTGAGTTTAGAGGGTGGACCGCTGGGTGCGCTAATGTCAAGTTCCGAGGGTGGACCGCTGC
 CTCGACCACTTGTGATGATTACAGAGGGGTCTAGAAATATGTACAGAAATGTACAGTTTATAAGAAATGTGACCTTTCCAGAAATGGCTGTGTC
 AGAAATCTTTGAACCAATTTCTGGAACAGACGTTTCTGGAGCATGGTCAAGCCATGATCTGTTGGCGGTGTCAATAGTACCTATCTCAATGGTC
 TTTTACCTCTCTTCCATCTCAAAAATTTTGAAGAAATTTTCTGGGTGCTGAGTTTAGAGGGTGGACCGCTGGCTGGCCCTAATGTCAAGTTCCGAG
 GGTGAGCGCTGGGTGCGACCAATGTGATGATACAGAGAGGGTCTAGAAATATGTACAGAAATAGTGTACAGTTATAAGAAATGTGACCTTTCCAG
 AATGGCTTGTCCAGAAATCTTTGAACCAATTTCTGGAACAGACGTTTCTGGAGCATGGTGGCCCGAGGCGCTCAGATTACCTGTTACACCCCTTCC
 TACTAGCACCTCTTCCACCTTCATCTCAAAAATTTTAAAAAATTTTCTGGCTCAGTGTAGAGGGTGGACCGCTGGCTGGCCCTAA
 TGCAAGTTTGAAGGGTGGGCGCTGGGTAGGCGCAATGTAGATACAGAGAGGGTCTAGAAATGTATACAGAAATAGTGTGCGGATTTATAAATTT
 AACTTTTCAACATGTGGCATAACCGTATGAACGTATAGCCAAATATGTACACAGTCAAAACATTTAGGTGAGAAAGATATAACATGTATAAATTT
 AATACTTTTATGATACATAATAAAAAATCATACAGTTTCCGACTCTCTCCACATCATTTATTAGGAATGGTATCTATTATACCTTACTTCT
 CTACGTTTCAAGATACATTGTTTCTCAGGCCTGCACTTTTCTTCAAGACACTTGGTTTCAACCAACATCTCTTAACTGTATGACGCTCAGGTTCTTA
 GGGAGTACGTTCTCTGATACAGGACACCGGTAATTTCTCCAGCTTGGCTACTCAACCAATCTGTCCCAACATTTTCTATCAATCAATGTTCCACA
 CGCTGTGCTCTGATGATTGGTACTCTGAATCTTCTAAACACTGGACATCTCGCTGTGCAACAGCACATTAGCTATCTTCTATGTTCTTCT
 AGAGCCCTTTAATTTCTTCTTATCCCCAGAAATATGTGTTAGCCCAAGGATAATTTGAGGATGCAATATACCCCAAGCGACAATTCAGTATTGGCC
 AAGAACAGGGTGTCTTATGATCAGACCGGTGGCAGCAACAGGGGTTCTTCTTCTTGGAGAGAGAAATTAAGCAGCTTCAATATGGTGGGTAT
 ACTAGGCTCGATCTCTCTGGGCAAACTGTGTTAATGCGGTATCATGAACTCAATATGCTGCTTCACTTAAAGTACATAGATTACCATAC
 AACATATTAATCTTGGCCCATGGGCTCCAACGTAGTGCATCAGAGACATCGTGTGTTGATTGGCGATGGCAATCATTAGGGGCGTTCACCATCGCTCG
 CCTTCTAGCATTCATATACCTTCTATCGAAATCGTATATTATTTCTGTATTTTGAAGAGCATCAAGAAATTCACGAGGTTGCAACAGCTGCA
 GCATGGAATACGTTTCTATTATTTTCTAATAAAGGAAAAGGCCAAGCTGTTGTTGTTGTAAGTTTCAATACCTACTCTTGTGGAAGGGGGAAGATA
 TTGACGCGCAAAACATTAGAGGTGACATGCTTTCGAGAGGCTTCAGAAATGTTAAGTTTTCTGTTGATTGATGTTGCTAGTATGGGAATATTGTT
 AATGATCAATAACCAATAGAGTATTTCCCTTTTCAATCAAAATAGTGTGTTGAGCCACGGATCTTGGTAAGAAAGCTCAAGATCTCCATGGA
 CACGAGACTTTTACACAAATCTAACATAGCCTTTAACTCATGAGGCTGTTATGATACAGCTTTTATGATACAGTCTTTTCAATAAGATTAT
 CGCACAGTTTGTGGCTCTTTTCTTACCATATCTCCCTCAACAGAGTCTTGAAGATTGTAATCAATCTCTTCAAGGGCCTACTGTTAATGATCTGCA
 CAACAGATTGCAACCAACCAACCACTCCACTGCTTCTGTTTCTTGAACATTTTGAAGATTGTTCCATCTTTCGCGAGAGAAATTTGGCATCTG
 TCAATATACGTAAAGGGTTTGAAGATGCTGGATTGTTTACGCTTCTCTATAAACACATCAACAACTCTACTGCCATTTACGCAACAGTTTGTGAT
 TAGTCTCAGAGTACTCTTACCATTTCTTCTTCTATCGAGGAAAGAAAGACTTGAATACACACTCTGTAGACACTAGCCCTTTCTGTTGGTCCCA
 ATTCATCTAGCTATAGATAGCGTCCATGTTCTTCTCTGAGGTCAGTTGACACGGATGATGGTGTGTTTCTAGGCAAGAAAGGCTCTCCCG
 ATAATAAATTTGCCATTGGATATCAGTGTGTTGCTTGTAAACACAGGAGATTGCTCCAAAAATCTGTTATCCGAAGATATGCAAGAGTTCAA
 GTGGTGCAGATTTTCTATTACGACCGTAATCAGAGGTGATATGACGATTCTGAAAAGAGCTGAATCTAATAACACTCGAACATTTTCAACGTA
 GAAACAAATACCACTCTTTCGAGAACTAGTAGACTTTTTCAGGCTGACCAAAACAGCCCTCAACTCTTGTATCTTCTATCAACCTTCTCTTCTTCT
 TCAGGCTGTTCTTGAAGTAACTTGAATCCAGTTCTGCTGTATCACCAGTGGCAAACTTGAATGCGCTGTCTACATACGAGTGTGAGTTTATCAT
 TAGACACAGAGAGATATTTTACGAACAAAAGTCTGTTGATGTTTCTGTAAGGCTTGAAGTTTTCGACGCTTCTTCTTATTTCTTCTTCTG
 AGTCATTTCCCTCGAGATGGGGGAAGAAACAGACTTGGTATTGTTGTTGATAAAGAACTAGATAATATCTCAAACTCTCTTATTTCTTCTTCTG
 AAACACTTGAAGAACAGGAGGAGAAATTTCTGGAGGTAATATGTCATTCAAGAGGCCAATTTCCCTTCTGTAAGAACCTCCAGAAATGATATATG
 GTTCAATGTTTCAAGTACTTTCAAGCACTGACGGTATCGTGGAGTCTGTTGACGCTGTTGATGATGTTTCTACATACGAGTGTGAGTTTATCAT
 GCGTACGCCGCTTTTATACAAAGATCCCGTGAAGAACTCCCTCGGTTCACTTCCAGGATAGGGGTGTGTCAGTTTATCATCAAAGTTAATATAT
 TTTTAAATATAACAAAAAATCGTACCGCTTATGGCTGCTATAAAGAGGAGGACCTGCTCACTTGGACATCATTAACCATCATATGAGGAGGA
 GAACATCAATATTGAACCTAGTCAGGAGATCTTGAAGAGGAGTGAAGAGGAGCATAGAATGGAACAGGAACCTTCTTGGACCCCAAA

FIGURE 2, Sheet 6 of 40

FIGURE 2, Sheet 7 of 40

FIGURE 2. Sheet 9 of 40

FIGURE 2, Sheet 10 of 40

FIGURE 2. Sheet 13 of 40

FIGURE 2, Sheet 14 of 40

ACGTGATACGTGAATGGGTTGGAATGGGTTATCACTCGCTTCTACTAGGCCGTTATGCGAATCTAAACACGGGAAACAGCGGTACAGTTTGTAGAGTTTC
 TCGGGGTGCGTTAAATGTAGGCTGTTCCCTCCAGTAATGCAGTTTGAACGGGGGTAATAATATCTCCACCTCGTACTTCAGTCAATGGACCACTGAA
 ACGTTTAAAGCACGAGGAGTAGAGATGTTGTGCGTTTACCAGGATATCTGTGGCGGGGAAAGAACACAGCGGTTATTCGAGGTTCTCTATTATATCAT
 TATCAGTTTTCAGATGATACATGGTCTTAGTTGGTAGGAGAAGAAATCTGCGCCGTTCTGTTGGTCCAGGCCAATCTGTCAGTTCGACGATCTACTCT
 AAAGTTTACGCTAGAACCGAAGCACCCGACTCTCGAAGGTGATTATGTACAACAACCTTTGTTCTGTTAGCATTACATCATGACTTATAACCCGAGAAC
 AATTATCAGATTGAGTAACCTTTGATGCTGGAATACAAATATTGTACCCAGAAGATGTGGTTAAACCGTCAGACGAGCTTTGAAGCCCAATCCAGC
 CCAACATTTTCCATTGCGCTTCTATTTCACTTTCAAGTACAGACACATATAAAGTCCGAGCTGGGTAGATGGATTCTAAGGTGGGTTATCTGT
 GCGGTTTAGAATCTTGAACCTCGGATTCATTTCCCTCCATATCTTTAACGTTTCATTGTATAGTACTTTGTTAATACCTGGTGTGAACAGTGT
 GTTTTCGAGTACAGAACGACACTGAAGGTTTGTAAATATCTGAAAGTTGATTTTCGGGAGGTAGTTAGCTGCTAAATGGACGCTATTTCAGCTTTGTC
 GGGAAAGACGGGAAATTAGCAGGGTAATCTGAACATATTTTTCATTCAATGTATCAGGAGTGATTTCCACATTATCATAGTTTATTGCTTTAGGCA
 ATCAAGCCCATGAGGGCTACATTTCTACAAGCAAAACCTCCGTTATTTTGGTATGGGTGAAGATATTGTTACCTGAGAAATCGAATACACGGTCTTG
 TTTTATCTCTGTTACAATGGAAATTTTCACTTATTTCCAGTGTGAAATACTTCCCTTGTGTATAGTGTATACCGTCAGCTAGAGAAATTTACTT
 CCACGTGTTGAATACGGCTGTTGTTTATTGCCCATTGTTGCGAGAGAACGCAATTCGGGTTCATTGGTATAGGAGCCCACTGCTGGGACAAATTCAGG
 GAAAGTTTTCGAATAGAGACGGACCACTATCGCTGCTGCCACCCCGCTTCTCTCTCATATTTTTCATCCACTAAATCAACTTTTATCTCTGCTTTAAAG
 GGAAGAGGACATTATCTGGTAATATTCCGTCAGAAATATGTATCTCTACGATGATCTTTTTCACCTTTTCTTAATAATTTCTCTGACGAGTATTA
 TCATTGCAATTTGGCAGCAATGATGATGATGATGAACCATCTCTATCAGGAGCCGATTGAGGAGAAATATCATCAACATATCATACCGATCTGCGGTGC
 AGTGGCTTTTATATCTGTGAACAATTTTGGATCTCTGACATCATGTCAGCAGTTTAAATCTCTCAGAGGATATCTTCAATATTTTAGGCGAGTTT
 TTGACCGCGGCTTCAACCATCTTCAAGTTCCTTTCTTTTTCGGGTCTTCTTAACAGCAGCGGAATCATCGATAGCCGCTGCTATATTCCCGAATA
 CAGCTTTTCTATTTTCGCGAGGTTTCTTATGTATCATCACAGATCTTTTCATCGGCGTAAATTTGTTCTGGACAATTTTCATACCAATATAGGAAAT
 AATAGAAATAAACACAGAGAGCGTGAAATTCAAAGGCACCCACAGGAGCAAAATACCCGTGACTGCTGTTTCTCAAACTTTGTTATGCTGTT
 CTTAAAGCGCAAGTTGGTTGTTGCGGAAACAAAAGTTACCGGCGATGGCAGGGAATAGAACCAGTCGTTATCGTCTGCTATGCTGATGCTAA
 ATGTATATCAGACGTTGAACAGGCAATGGAGTGTGTTGCGCAGACAGCAGGATGATTAATGACTCGCTAGCCAACTTAAATTTGGGCGATTCTCTT
 AAAGAACTGATTTAATTTGGAATACTTGAGATACCGCTCTACGCCCCCTCTTGGGGAATTAACACGACAAACAATATGCGGCAACAGTTGACA
 TCAACCTAATGGCTCATTTCTCTACGCTGCTTTGGGTATAGAAAGTATATCTGAATTTATACGGAGAGTTGATGCTGAATCAACGTAAGAAATA
 TGGAAAAAACCTTCTGAACCAATCTCACGCCCTCACCGCTGGGAGGGGTAGAACCTCTCTATCTGTCAGAGTTGGCAAAATGCAATGAGGCAAGTTT
 ATCAGCATGGGGCGCTTGGACAGATTGAATTCAGCAATAGTGACAGCGGCTTGGGGCTATTTGCCAGTGAACGTAACATTTCTACGTTGAAATCTG
 TAACTACATCTACGATGTAGAATTTGCAGAAAGAGATGCTGCTACTACAGATACAGGGAATGTAGTCTATCTTTCCACCAAAATGGACGAGATGAAGA
 TGACATAATAACCGTTTCAGAAATATTAGATAAGGTATCAAAACGACCGCAAGGAAGGTATAGACTGGCGCCCCACCCCTGACAAATTCGTTCCCTTAC
 CAATTTGATTTGGGCGCATGATTCTGTAGATGACTGTTCTTATAGATCTCATCAACCATCTGATCGCTGCTGAATTTTATGGCAAAATTTATCTGCT
 TCATATGTAACCATTTAAGGGCAGTTATAGGAGTATGAGGGAATTTTATACGTAACATTTCTTCTTATCCGATATTTATTTGAGGATGGACGTA
 ATGGTCTTCTGTTGAACCTGTACAATAGAGTGGATGGTGTAGTAGTTAGTTTGAATTTTCTCCACTCAAAAGAGGATCCTTTTCAGGA
 GCTGACAAATGTTAAGCTGAAAAGACTTCTGGTGGTGTGGAGAGTTTTCCTCCGCTTCTTGGACACTGAATGGGCTCAAGCTAATATAACGTCAT
 GGCTGTTTATTAATAACCAATATAATAGTACATCCCTGTGACAGAGACACCTTAATGAGACTAGCGATAGGACGAGTACGCGTGGCCGACATCC
 TATTTTCGACGAATTAACCTCTTGACAAACAGCAGTGACCAACCGTATACCTTCCAGTCTGCGAGAATTTGCTCGGGCGAGCTCTG
 GACGAAGAAGAGCTGGAACAAAATGCTAGTAAATCAGTCAAAGAGACGGGAGAAGAAAGGATAAGAAACAATACGTTCTCTTCTTATGTTTATTAC
 TGAAGAACACAAAATAGAAATTTGGAATTAACATAGGCGATACAGATGATGAGACTACAGATGTGGCTTCTTGGGCAGCTACTTCTCGACATCCCT
 TATCCGTAAATAGGACATATCGCTTTAAAAAATATGGGGCTTGGAGGCTTGGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT
 GTCACCGATAAGAGCAGTCTCTCTTATTTCCGTATGTGTCGAGTGGAGTGTCTTACTATTACATCCCTGTTGTAAAGCACCGGCCATAATTAAGTG
 TGTGGTTACAATCTGAAAGATTTTTCCAGGAAATATAAAATATAAATGAAAGGTACAATCTCTTCTCATGTAGATTTGTGCAAGAACTCAACG
 CGTTTAAAAATAAAATTTGCTGCCGAACGTTCCGAGGTGAAAGGATTTAAATACGTAAGCAACAGGAGAGAAAGACGACATCTATACG
 GAACATGTATTTGGTTAACGATTTTGTGGAACCAAGTCTGTCAGAACACTCTCAACCTTTTGGAGAATTTTCCGTATAAAGGGGAAAGTACTCAACA
 TGTTCACCTGATTGACCCGAGCTTCCGCAAGATGAACAACCTAGTCAATCCGAGCTTTCATGACGTTTACCAGGTGGTGGCCGAGTTATCTCTATCAGA
 ATTGGAAGAGGATGTAAAGAAATCCGGAATCATCCATATATAGAACTCCGATATCTCTTCCAAACAAGGATATTGTACATATTGCTGATTTACG
 CTCTCTCCGAAGACGCACTCATTTCAAGTTCTATACCCAACTCAAGAGGTCATCGAACACTTCCAGTAATCTTCCACTGCAACCAATATGTCCTCT
 TCTGGGTACACCTTCTGGAGCAAGCCATCATCGCTGCTCCAGAGCTGCTGACGTACGAGATTGTGAATGCCAAGTACAGGGGATTTGTTGTCATCCC
 ATACTACAGCGCCCATCAACTATCAAACTGGGAAGAGTCTACTGATGAGCAAACTGGCGTCCGTAAGTCTGACATTTTAAATGAGATGTGGATCA
 TACAATTCATCTTAAATGTGATGATCAACAAGAAAGAACACCAACTTCTTCTACTGCTGTGCAAGTAAATGGGAGAACTTGAAGAACAGATGA
 TGCTCCACATCTGCTGATTAATGTTCTTTGCCAACCTTACTACTAGCAACCTTCCGAGCGTAGTAGTTTCCCTGATGCTGACAGAGAGCAAGGG
 GAAACACCTGCCATCTAGCAATCCAAAGAGATATGCTGATGCACTCTGTTCTGATCTCCCTCTACGGCGCACCTGGTTTCAAGATAACAACCTCG
 TACATGAAATGCTGCCCTGAACTCAAGTCTAACAAGTGTGCAAGTACTTCTTGCAGCTGACAAGTACGAGATTTTACCCAACTTAAACAACAAATC
 AACTAGAACAGATACCATGTGTGGAGTGTGTGCAACATCTGTGGAAGAGTGAAGAAATGAAGGAAACAACAGCTTTCTTGGTATACAGATTAATG
 CAAGCATTAATCCATTGCGAATGCCTGATGGGAATGTGTGCTGCTGCTGCAATGATCAATGCCCATGTGCGGTGAGGATGTGGGCGACGAGTACTG
 GAAAGATGCCCTCTACAATATTAGATGGTTAAACTGGCTGAGATCTGAACACATCTGTGACTTTTGAAGCAAAAGCAAGATTTCTAAGC
 AGATGGAAGCAATGAACCTCCAGAGTGTGTTGTTCTCTCTGCGAGGACATTTCTACCCAGGCCAGAGAGCGGCAAGCCATCAAGATGAGAGAGA
 AATTGCCACCAACGCCATCGCTGAAGCCACAGCTCAAGGAGATGTCAACTCTACTTCCCTGTTCTCATTTGACGGGAGCGGAGAGAAATATGAAGAGAG
 GGAGAAGAAATTTCAATTTCTGAAGAGGAGCGCTTGCTTTGGAAGACATTTCTGGAAGATGAGGAAGAGCGGACCAATACAGATGCGCCAGTTTG
 CTGAAGTGTCTAGAGAGGCGTTTCTGTCAATATTATTAACAATGATAATCTCATCGACACATCTCTACAGTAATATTGTGAACAGTTTATGGAGT
 TGAAGAGTCACTGCTGCTTCTTCACTATCAACATGCTCAAGAAATGACGCTTTGAGTCTATACGCTCAAGAGATCTCGAGTTGGAGGAGAAAGAGTG
 CCCGTCTGAACCTGTCCAATGACAAGAGGCGATTATTCACAGCAGCTTCTCCATGCTTTTGTGACTTTGCCACAGAAACAACCTCTCAAAATGTTGGAT
 TGGACTTTCAAGCAGTGTATGATCCCATCATATCCAACTATATCGAGAGCTTTGGTAGTCTTCTACGCCATCCAGGAGCGGTCACTTTCTGGA
 CGGGGCCAGGACTATTATGCGAGAGTATCAGATACGCAATGATATTGTCTATTCTCAGAAATGGCAAGTGAAGTGCACATCAGGAGCATTAGAT
 GTCTTTGAGGGTAGTTTATTATCCCACTGTTCAAGAAATCAGGACTGGAATCTTACTCTAACTGGAACGCCATTGGAGCGTAGAAATTTATGCTC
 GAGATATTGCTGAGGAATTTGTTAGGGTATGTGAAACCTCTCTAGCTTCAAGCGCAACCCCTGTTCTATGATACCTTTAGAGATGGAGCAATCCC
 CATTTCTAATGAATATAGTAGATTTCATCCACCACTCATCACTGGTCTATGCAAGTTAATGCACTCCATTGTATGAGAAAGTACATTGAACAGAG
 AATACAATGTGCACTGTAACTTGGCTCTACTGATGAAAGGTGGAGTTTAAAGGATATCTCAACTCAGATGGAGCGCTTGTTCATGAACAAAT
 ACAACATAGAAATGCTCCAGCAACAAAGATTGAGCTCTAGGAAGTCTTCAACCATGATTTGGGTGTGCTTAAATTTGGTGTATACAACTCTGATA
 TATTATGAAGTACTGTTTACTTTAATCTAAACAATAAAATAATGTAAGAAATTTTCTATCTCCCTCTGCTGCTGATGATGATGATGATGATGATGAT
 GCGGGGTAGCCTCTCCAGGACACTTGTCCAGTAGCGTGTGCGAGTGAACATCATGGCAGCAGACCTCTAGAGTTGGCTATCCAGGAACAAATCC
 AGTCTGAATTTGAAGAAATTTCCGATGACTGAATTTCTCAATTTCTTCCCATAAAAAGTGGCATCTGCGAAGAAGTGCAGCTAATGGACGGCCATCT
 TCTACACTAGAAATGAGGAACGAAGTTGACCAATTTCTGTTGCCAGTAACAGGAGCTGAAACTCTTGGGGCACTTTTGGGCACTTTGTATGGAG
 GCATTTATAGCTGGTTCTATAGATGCTGAAACGTCGCTGGGCTTTTGGAGTGCAGCAACTGGACTAGGATACCTCTATTGAAAAACTAGCCCTGA
 TTGCCGCTGAGGATAAATAACAATAACAATTTGAT
 TTCATCAATACAGAACACTTCTCACAAAAATCTTCCCGGTGACTTTGAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
 TCCAGAGGAGCGGCAAGGAAATTTGAGCAGCAGGAAAGGTTTGTCTCCAGAACCTGATTTGTTGCCATATTAAATACCTGCATACAGCAAGAACCTTTG
 TCTTGTATACAGGCAAGCACTAGAACAGCAAGTATTGGAT
 ATTGACTACTTTGCACAATTTATTTGTTAGGAAAGTCTTCCCATCATTTGTACAAATGCACCTTTCCGCTCTTCTCCCTTTGGACAACACCTATCATC
 AACATTGAAATTTCTCATTTCTTAAATGAAGATACGACCACTATTCAGCGTCAATTTCTATACCATCACTATGGTAAATAAACCACTACCGGAAAAA
 ATAGTAGATGGAGATGCCCAATAACTTGTGACTGCGAGCAAGCGTCTATTTTACGTGGAGTTTAAACAGTCACTGGAGATTACGGATGGTTC

FIGURE 2. Sheet 17 of 40

FIGURE 2, Sheet 18 of 40

FIGURE 2. Sheet 20 of 40

FIGURE 2, Sheet 21 of 40

FIGURE 2. Sheet 22 of 40

TTTCCATACACATGATTCTCCACATGTTATTATTGTTAGGTGGGAATTTCTTAAACGTAACCTGGGTTATACGTGGAGCATTGTTGTTGTTGTTACTAC
 TGCTCCTGGTTAGGGTGGTTTCTTGTGTTTCATCTCCTTCAATTCCTGGATACACTTCTGGATAGATGGAGATTGTTGAGGTACTTCTGGCAGCTATTTC
 GTAGACATGTTCTTCCACATTATCATCTCTCCTCCTCATTCTTGATATTGAAGATGGGAGAAGGTGGGTTTGAAGGAAGCTTCAGATGTCGTCGGTTT
 ACATATGACACGTAATCACCGCCATTGCTTCTAGGTGGAGATGGCGCGGTGTTGGCAGAGGAAGACTATTGTCAGAGAAGAAATGACGACGACGCTCG
 ACATTGGCATATAGACACTTCTTCTCTTTATTGTTACTTGTGGTGTGTTGAGGTGATTGGCACTGGCCTGGGTTAGGCGGAGCTAGCATAGGT
 TGGACCAGAAAATGTTACCAGTTTGTGACTTGGAGGGTAACATTGGCCAGATCGTAAGTGGTAGATTGAGCATGGGTTGAGAGATGTCAGGTCCTTA
 TCTGGAGTGTGAAACCGGACGTGGCTGATGAACTGGATTGCTGCTCCGTTCTGTTCTTCTTATTCTTCTTCTCCACGGTGAGCAAGAAA
 TGAAGTAGTAGTTCTGTCGATGCATTCACGAGCATGTTTCATGGCCAGTGGCAGCCCAATAGTCATGGAGCAGCCGACTGCCAATCCACTCAGGAAAGT
 GACGAGGATAATGTTGTCGTTGGTGATAAGATTGCCATCTCGAGAGATGTTCTTCTCGACTACCGAGAGATATGATGCTCTGGGGCCCTGCTTCG
 TCCTTTTATACCGCTCACTTGGTTAGAAAAACGATAGTGTAGTTACAGTATTGTCCTTCTGGTCTTGGGATTGGGGTTCTGATACTCTGCCAGGCT
 CCAGATATAGGTGGTAGACCGAGATTCTTCCCTTGATGCACTTCTATTCCATCCAAGATATCTTACCAATAAGTCTTATGCAAGGACACCAGAAT
 ATTGAACCTCGTGAACGCTTGACTTGTGATGATGCATAATTCACCTTGTTCGACTGTTATACATACTCCTTCAATAGAATCGTATAGCCCGGGTTATTG
 GCAGGTCCTCTTACGGTAGAAAACTGCAAAATATTGTTAGGCACAACTCTGGCAGCTTACAGAAATGACAGTAATGTTTATCTATAATGTCACCAG
 AAAAGTGAAAATCTATCTCCGACACTGGAGAAAGGTTGAAGGAAAATCTTCCGCAATATCTTGAGTTCCCGCAATACTTCTGTTCTGAGATTGTTGA
 CTTTTTAGGATTATTGTTGGCCAGACTTACAGCATCGAAAATCTTCTCCATCTCTTCCAGCGGATTTTCCATCTTACTTTGTTAGTGGTTTTCAGTT
 CATTAAATGTTGTTTACGGGAGGCTATAAAACAAAATCTTACGCATCTGTTTCACTTCACTAGTGGACGAGCTCGAGATGACGAGACGTTGGTGTCTGT
 TCCAAAAGCGGCTTCTAGGATGTTATCTTAGGAAATGTCGACTACACTTCTGTTACTACTGCAATAATGTTGTCAGTCTAGACATCGATTCTCAAGGAC
 AATATCACAGGCAAAAATCTAGTTATTGAACAGAAATGGGTGAAGAAAACAGAAAATAAAGAAAGGAAGTGCACCTGAAACAAAGGAAATTA
 GTGACGAAGACATATACGCCACCAAGGAATTCGAACAGACAATAAAGGTCCTACAGACAAAAGGTCGCCACGAGGAAACGCCATCGCGCGCCGACG
 TGCCCGTGCCACTGCTGCTGGGTAGAAAAGGCTATGCTATCAGAAAGTGAAGGAAAATCAATGGTCATCAACAGAGCTAGAATGGTGTCTTCTAAGCGA
 GACACGTCGCCAGAAACGTTCTCGCATTTGAAGAACAGGGAATCTTCTTCACTTGTGATATTGAAACTGGGATCAGTAGTGTCTGCGGCTTCAAG
 ATCCTTCTGCTTACAAAATGTTGTTGATTAAAGCCACGACTGATATGCTGATATTCTACAGAAAACATCAGTGTGGCTAACGTGTCTATAGTGAATA
 AGTGTCCACTTTTAAATAGATTCCACTTGAACCTTATTCGACTCGGAAATCTTTCGAAAGAAATGCTACTCTTACAGTTATAACCCAGAAACGTTCCCC
 GGTATGTTTTTCAAGCTGCGAGTGCCCGCAAGCCCTCTTTCGCTGGAGAGACTATAGGGGAATACTACAAAGGTTGCAATGATGCGCGATAGTAAAG
 ATCCCAATTTTAAATGTCGACTGGTTAGGATAAAACCTGCAATTAACATTTAAAGTTGGGAAAATTAAGTGTCTCGGAGAGGAGAGTGGTGGCG
 TGATGTTTCTGTCGTATCCAAATTAATTTTGGTTTATTCCTATTCTTATGGAACAACATTAATGAAATGTCGCCCAAGAGACAAAGAGTCAAGAA
 AAATACGGCATCCCGCATCTAGAATGCTACTTGTACATTGACATGTTGCTCCACTCTACCCGTACGTCAAACCATCGGCCGAGCAAGTGAAGAGGCGA
 TGGTGGACCAACAACATATTTCTGAAGTGGATAGGACATACTATTGAACAAAGAACAGTATGGACGCTGCCATGTCGCAAAATTTAGTGCCTTCAAAAGA
 AGAGAGTATCTCTCTTCAATTAATAAATAAGATCAACAACCTTTTGGACATTTGTTGAACCTTCAAAAGAACTACTCGACGTGCTATAGACACACTT
 TCTTTCGATCTATAAACCAAGACAGGTGGTGAATAAAATGACCAATATTACGGTAAAGAGAGATGTACCCGTTTCTGTTGACGTTTATGTTCTGTG
 TTTCTGAATAACAAACAGTATGATGAATAGTCGATTTCTTGTCAAGGAAATGGTGGCTAGATGAACCAAGTAAACAGGATAAATGATCATATTGT
 GGATTGTTGTACAGAAGAAATAGTGGAGGAATGTAATCAAGGGTTTATGGCTGCCCTTTTGGAGGAGCACCAGAAGGAAAATATCAACGCCCT
 TATGTTTTTATAGCGAGGCTGTAAATCAAAAAATGGTAACAAATAGATTAACAAATAATAGTAACTATTGTCGGGTCAAGTAGGGCGAAGAGGA
 ATGCAAAACTACAGGAAAACACCGGTGTAACCTTAGCCAGGTGGAACAGATGATGGGCTGCTACCGATTTTGAATAATACATCTCAACAGACATTGC
 ACCGTGATTTTGCCAAAGTTATTGGTAATGATGTATATAGTTTATTACATTTAATGACAACTTGCCTAAATCCCGTGGACATGCTCTAACATACAAAGAA
 AGGGCCCTTTCAAGTAATGAAAGTACATATAAACACCTGGAATGCAATTTAGTACTCTATTGAAAATCCATTATAAATAACCAAGAACTGCTA
 ATAAAGTAAACAAATCAAAACCTAAATTTCTCGAATCGGACAGAAAGAGCTCTTTCTGTGCAACGCGTGTGGTGTCTCAATTTGAACAAGGTAGTGA
 TGAATCATAAAGGATTTGTACAAGTTGCGATCAAAATACTACAGTTTACATAGAGAATGCATTATCTGACATTAACAGAGACAAGAGATTAAACGT
 TTTAAAGCAGCTGCAACCCATCCGCCAGTGAAGCAAGAAATGGTAGATTCTTATCTCCTCTTCTCATCTCTTCTTCTTCTCTCTGACAGCTCTAACA
 AGAAGATAGATGCACCCCTAGTGATTTTATAGATTATGTGACAAATTTCTGACTGATTGTTCTAAGAAATGCTTCCATGACTGGGTGATCTAAGACAGATT
 GTAGGTGTTTGTATGACTACACTGTGAGACGGACAGAGATTGTCGCTCTCTTCCGAGAACACGTAATCTGTTGACGGAAGGAGGTTACAAAATTTCTTCTG
 GTACCAACAGCTCGGGAAGATGTGCCAGGAGTTTATAGTGCGTCTTCTTACGAAGAGACGCAATGATCTTGCCCACTATAGTATGATCTAAGAA
 AATCCAAAATTTGAGGAAGAACTCCCTGAATCATTTTGTGGTTTATCAATCAAAATGGAGGAGATTGTTTGTGAATAAGCGAGCCGCTACTACGACA
 CGATGCATCTAAGCATAGGGAACCTGGATAATGTGGACACTTTTGGCCAGGGTTTATAGTAAACGGATGGCCCTCATCTTGTAGAGAGGACCTACTGAGGAA
 TTGGCTCTATACTTTTCAAAATGATAAGATTAAATATGAAGAGGCAAGAAATGGAATTTGGAATATAACACAGGAGGCTGGCACCAGAGGACAAAT
 AAAGAAAGAAAGATGCGAAAAAGAGGATCAATCTCTTAGCGTTAGTGAAATTTGGATGTTTAAACGGGCACACATGACCTATGCCCCGAGGGCTA
 GAGGTTTATCCAGAAAATATATCTTTGTCAAGAAACGAATGAGAGATTAGCCCTTAAGGAACCTTTCCCTGAAAGAACTATCTCTCAGGT
 TTTGAGTAGGCAACTGATGTATCTAGCGGTGAAGATTATGCAATGAAGATGATAGTCAAGGAGGCAAGAAATCCATTTTAGGCAACCTGATTCTGGA
 GAGTACGTGGCTACTTGTGATGCTTTACTCGGAATATTAAACAGGCGCTGCGTGAAGCACAACATACAGGTATGTTATAGACTACGACAAATGGA
 AAAGGCTGGAAGCTGAATTTCTAAGTACTCTGACTTCAATTTTAAAGAGCAAGAGCTGTGTGTAATTCGACAAATCCAAATCTGAGGGCAATTTA
 TAGTCCAGATAATAAAGGTTTCTTGTGTCGCTGTAGCTGAACCTGTAAAGACGGCATTAACCTTTTAGGGGTTTACACGAAACCGCTCTGATCTCGAG
 AGAGATATAAATCAAGCTGAAAATCTACCATCAATTCATTTGGTGTAACCTGGCCCTATGTGAATCTCTTAAATCGCATCAAGACGATGACAGTAAAT
 TTGATAAAAAAAGATCGGGGAAAATGCTGCACTTTAATATTGGACAGAACTCAAGAAAGCTGCTTACGCGCTTCTGGATGTAGCTGACTCATTTGA
 GAAAATCAAGGGAGAAATCCAATCACCTGAGGAGGCTGCGGCTCTTCTGTTGCTCTCTATGGAGCACCTCCAAACCTTCAGCTTGGCTGTGGCCTCT
 ATCATCACTGGAGAAAGAACATCTTAAACGACAAAATATCTATCGGATAATGTCCTATTGAAAATGCTGTTGCTCGCGTTGGACAGAAAATATCGCA
 AGAGAGCCGACCCAGGCACTGATGAAATTAGAACCATCATGGAAGATATTACAGGAGTTTGTCCGGTGGCTACAGGCAATATAGCCCGCTCGAGGAAGA
 AATAAGGTGCATATAGGCATCATGAATAACAAACGCCATGACTGTTTGTGGATATTATACAATGGACACATCTATTCTTCCGACCTCTTCTCTA
 ACAGATTTCAAAACCCCACTGTCTATTGCCAATGTGACTAAGCGGATGGAGAGCATTTTTCAAAGGTCGACTCTGCTAGGTCTACAGATTTCGACGCTT
 TTGTTAATGGTGTGCGAATAATATGGATATAAAGTCATCAATAGATTGGGCAAAATGGTAGAAAATGATGATCAAAATACAGATTCTACACCTACACCC
 TTGTTCAAGTTGACACTATTGTTCCAGAGACGCAAGTGTAGTTAAACAGCAGCATTAATGATATACGCTTCTGTTGGAATACTTTATGTCGCTGCTGCA
 ACACAGCTAACCTTTATAGCGGAGATTGAAAACCTGCAAGAGCTGAGTTGTTTGGAGCACTCATGTCGCGTATGAGGAGAGGCAATCTGTTGTTG
 CTACTTTGAGGAAATTTAAGCGAAAGAGCCCAATGCAAGTGAAGAAATTTGGTGTAAACGATACAGGATAGATTGCTAGATGCACTAGCAATCTCAC
 CGGACCTGTGATGTTCTCATTTTAGATCTTGTACGTACTTAAGTGGCAGAGATCGTGTCAACTAAATTTCTCCGTTGAATTAAGGAAACCTACTCTT
 TGTGGAATAGGTTTGTGTATAATAAAGGCTCAGACATCAAGATGGGAATTAAGAGAGGTTTGGAGGGAAGTGGAGGAAGTGAAGGCAACACA
 ACCCTTTCACAAAGAAATCTAAATCGGAACACCCAGTTTGAAGCAAGAGTTTATCCACATCTTCTAAGAAAGAGCAAAATCCAATAACACACCAAG
 ACCAAGGAAGCAAACTTCTAGAATTCTGAAGGATCTGGAGCGGAGCCCACTGCTTCTGATGAGAAGGTCAAGCAAGAAATGAAGAAAGTCCCT
 CTGAAGCTATTGCTGAATTTTTCATGTTTGGGATCGCTCAAGACAGCAAGTTCAAGAGCCTTCTCCCATTAAGCCGATCAAGAGCATCACTACTA
 AATTGTTATCGATGCAATTAATCAGCCTGTGCGCAAGATGTTGGTGTACCACTTATCACTTTAAGGAGATGCAAGATGTTGTTGGAAGAAATTAAGGAC
 GATAGCGACGAAACAACTAGCGCTCATTTAAGAGTAAGAAATCCCCCAAGAAATTTGACCTCTCTTCCGATTACGTTGATCGCTTAAACAGGATTTC
 TGGTGGTGTAAATTAAGAGGTTGGCGGAGCTATTGAAGTAAGGAATTTGTTGAGAGTAACAGCATGATCATGAACAGTGTCTGGGTACTGTTGTGTC
 CAACATTCCTTACACATGAAGATTAATATTGTTGTTTGGTAACTTTATTGTTACATTGCTAATGATGATTGTACACATCTTTAGGGATGAT
 GAGAAATTTGAATGAGTCAGGTAACAAGATCAATTTCAAGAGGATTAGAAAATAAAGAGATGGTATAAATTAATTTTATTCAACAAAACCTT

FIGURE 2, Sheet 24 of 40

FIGURE 2. Sheet 25 of 40

FIGURE 2, Sheet 27 of 40

FIGURE 2, Sheet 28 of 40

FIGURE 2. Sheet 29 of 40

FIGURE 2, Sheet 30 of 40

FIGURE 2, Sheet 32 of 40

FIGURE 2. Sheet 34 of 40

FIGURE 2. Sheet 35 of 40

FIGURE 2, Sheet 36 of 40

FIGURE 2. Sheet 37 of 40

FIGURE 2. Sheet 39 of 40

GAGGAATTCGAGGCCCTATFGGTGCTCCTGGTAAGCCTGGCACGGAAGGGGTTAGAGGTCTAGAGGGGTGAGAGGTGTTCTGGCTATCCTGGCGCACA
AGGGGAATTAGGTCCCCAAGGACCAACAGGTCTCAAGGGCCAGCAGGTCTCAAGGGCCGATGGGGCGTACAGGAGATACTGGTCCCATGGGCCCTCCT
GGAGCAGTGGGACCAAGAGGAGAGAAAGGAGGTAGAGGAAGAAAGGGAAAAATGGCCCTAAAGGAGCGGACGGAAGAATGCGCTAAATATCATACAAA
AATATTCATCACCCTGCTCCTGCAGAGATAATGTGGGAAGGAAATGAAATCGGAGAAGCATACATTGGAAGATCTTATGGAACGATACAATCCCTGT
GATGATAGAAAATAGAAATAGGGATGACAAATGAGGACAAAAAAACGAATATTGTATACAAGTAATGACAATGCACTCAATAACAACTAGAGGAAGAACA
TCGGGTGTTTTGTGGTAAGCAATAAGACAGATTATATCCTTTTAGTTACTTTACTGATGCCAGAAAGTGTTCCTGTAGAACAGATGTCAGTACAAATG
CGAGGTCAGAGAGGGTGAATGCTGTTAGAGAAAGAGAAAGCAAATCGTACAGATTTATTAGGCCGTCTGACCAATCTATAGGTACTCATTACGTTCAAA
AATTGCCGTGGTAATGTATCCAGACGCAAGCATGAGTTACTCAGTTGATACATTAGACGCTGATGTGGCGCGAAGAGAAACACGCTGTGCTTTTATTA
GCAGAAACCATACAGGGGAAAAAGATAGAGGTTTCTATGCTGATAGAGGAAGTGTAGGAGGTTGATGGTACCTCCCACTGAAGAAGAGTTATTGGTAT
TGCAAGC
(SEQ ID NO:1)

WO 01/38351

42/201

PCT/US00/28888

CT1035

Nucleotide

Genomic coordinates:

Start: 130589

Stop: 131444 (SEQ ID NO: 2)

Amino Acid

MAAAKMDAILADINGNDTDLKSLITDVIQKRAKAVMDRNRKMDMNRVDEAIQEAVA
 KQKALVVFQKLVEETDSGQSVPTLSGSDYDAWVDRAMPSHIELVESVEGDSLYDKLP
 NVQDIDDQIGDEIDTPISYLAMVVVKVDCETGDIEEYNLAPTFGVTQNNKIYRDERDQ
 FTKADKSVRIFKLAKLDSISGKSRQLTYAVKNNNEYTEFVCSVFAEFESDSDTTKSGIG
 REYDKPKNEFEYEEREIFTFFIPIQPAGTKLLLYFLVDVRSRII
 (SEQ ID NO: 3)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q65326 (Q65326) DNA POLYMERASE (FRAGMENT) | 36 | 0.31 |
| MYSB_CAEEL (P02566) MYOSIN HEAVY CHAIN B (MHC B) | 34 | 1.6 |
| O02244 (O02244) UNC-54 PROTEIN | 34 | 1.6 |
| O04327 (O04327) CELL DIVISION PROTEIN FTSH ISOLOG | 33 | 2.1 |
| Q04009 (Q04009) MYOSIN HEAVY CHAIN | 33 | 2.1 |
| O02077 (O02077) COSMID C48E7 | 33 | 2.7 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 765 of CT1035: this
 corresponds to nucleotides 130649 to 131413 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=350

5'stop=373

3'start=405

3'stop=427

5'primer=TCCCTCCTTTTAACTACAAGACA (residues 350 to 373 of SEQ ID NO: 2)

Tm5=57.95

3'primer=CCATGGCAAGGTAAGATATTGGT (residues 405 to 427 of SEQ ID NO: 2)

Tm3=57.97

probel=CGGTGATGAGATAGATAC (residues 387 to 404 of SEQ ID NO: 2)

probelstart=387

probelstop=404

direction1=Forward

Tm1=69.08

score1=1.91

length=78

CT1037
 Nucleotide
 Genomic coordinates:
 Start: 131480
 Stop: 132941 (SEQ ID NO:4)

Amino Acid
 MFVISIATSLVLFFLLFVSITILDGAKTIDSQPFRRKRKRKYRTSESGDGIDGGTGTT
 NGGGGGGGEGGGGGTNGNGTGTTNGGGGGGEGGGGGTNGNGSGTTNGGGGGGEGGGGGTNG
 GGGNGNGGGNGNGNGNGGDDTDDDFEPTALLKERLLNSISSKPKEYEAFVSAEVETAL
 QLSRDDSTQTIIIDDDQLELDASDTLQGKPRDYLFKLAGVSSAFLEGTTIRKAEDRARNI
 NEEIEIAQTILSQLREKHINDEYDGKYATPEERADFSNSLNLVTKYTNHEVGLLVGETIEK
 AFPHEIEFERCIILVEDFNSGTITSNTMQYRSNAYKIRVVEGSTTDPGEVVPDDCLVFAV
 VVNKEQHSLEISATNRCQDICFVIIPRLSAIGKNATMVIRKGDEIKQETYLFVANKNDTT
 HFSIITDKDES VGIENMLIFSERILPTLSDPATVPRPLTDANVLSAYGKRLGVGAFTDK
 NLLSSQ
 (SEQ ID NO:5)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| 065514 (O65514) PUTATIVE GLYCINE-RICH CELL WALL PROTEIN | 120 | 3e-26 |
| Q21835 (Q21835) R08B4.1 PROTEIN | 112 | 6e-24 |
| Q43522 (Q43522) TFM5 GENE | 110 | 2e-23 |
| O53553 (O53553) PGRS-FAMILY PROTEIN | 108 | 1e-22 |
| O65450 (O65450) GLYCINE-RICH PROTEIN | 107 | 2e-22 |
| GRP1_PHAVU (P10495) GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN | 106 | 3e-22 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has sequence identity to nucleotides 392 to 1365 of CT1037: this corresponds to nucleotides 131892 to 132865 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=723
 5'stop=745
 3'start=792
 3'stop=811
 5'primer=TGAGGAAGAAATTGCACAAACAA (residues 723 to 745 of SEQ ID NO: 4)
 Tm5=58.39
 3'primer=CCTCGGGTGTGGCATATTTT (residues 792 to 811 of SEQ ID NO: 4)
 Tm3=58.37
 probe1=CAACGATGAATACGATGG (residues 774 to 791 of SEQ ID NO: 4)
 probe1start=774
 probe1stop=791
 direction1=Reverse
 Tm1=68.98
 score1=1.98
 length=89

WO 01/38351

44/201

PCT/US00/28888

CT1038

Nucleotide

Genomic coordinates:

Start: 132993

Stop: 133896 (SEQ ID NO: 6)

Amino Acid

MGDKQKVEQLLRELKAEANDDWLSANVDPIVERFVTTKSDETAQVVKQAVDEKYDELLED
 KVEEMRPDIINEASETYDKLAADMIREVDTSSVIAPAIAGTVARTINNLRDKRKEYEKRL
 WTLAYKPWRRYVQAITVMEFRLSYKDLTVHANSPTYLTFPFLRIKKIAYINNDRASPVNC
 SLSVSYPNKSEWGNDNGVGRKVDIHIRRNDLQEKDLYLSVICMLDTEFSGYDKAVEVDAH
 KFHFEAGNRTMFLPKTSNLFNRSHIVNSKICTIVFPASASSASTTELDNVYYRITCTCS
 (SEQ ID NO: 7)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| RESA_PLAFF (Q26005) RING-INFECTED ERYTHROCYTE SURFACE ANTIG | 38 | 0.12 |
| Q9X9G6 (Q9X9G6) HYPOTHETICAL 35.5 KD PROTEIN | 37 | 0.15 |
| CAA21397 (CAA21397) ORF 74 | 37 | 0.15 |
| Q9ZfZ9 (Q9ZfZ9) HYPOTHETICAL 21.1 KD PROTEIN | 37 | 0.15 |
| RESA_PLAFF (P13830) RING-INFECTED ERYTHROCYTE SURFACE ANTIG | 36 | 0.26 |
| Q9ZH03 (Q9ZH03) LAMBDA HOST SPECIFICITY PROTEIN J | 36 | 0.45 |

CT1039
 Nucleotide
 Genomic coordinates:
 Start: 133968
 Stop: 136344 (SEQ ID NO: 8)

Amino Acid
 MEEESQVRVQRRIGVLPEEAASQILKDTKLRSYLGVGHWGYSVSVIKSALQKGCRRNDED
 ITAWSIREAYLYYHLGLNYIENVKPAAKSLNTNMVNRIKIIAEDTSPRSMVASNECVRT
 LEKYEKGNFRQPSYLMDAAMRLVHASSSRVCSHMRALCCKEEDSDKLGGIYYANFNELET
 QCVSAVNFSPIERIKHVFREIESVKLGKKSQVLLNLSVAAYHVLRYYGDKVKDTNKKHS
 GPFKRKEFEQFWGLCFKFVTQHVKTDPELRICYFNELTYAINWRRDFFCSKGFFREESLFL
 TSIVELIIAMCIGDRKQFAKIQKRD LKRFNKGEGRKEEAATFDWIEGHVKRMPQMPVWV
 LDKHTNKNTHGVSFALESSMVSGGDKRWSPGVWLHSYTKMRDLDSPPPPEVGQFLDQAFNT
 LKREAASHCVTRNICTTTGFIKASSFTANINSEPMIEKEIKKRKIEIKDDNTTATVTVS
 ATTSSSITSTPPPTKKQKTPSGSNKVDSIQLNNLPTLNMEDLDREVHVNQNSKKGVA
 TVLMKDGNKVVFEMRKSFSGWGSQNFVQVLKDEDVCKLDYLLPCPDSPYRGLYRCYFK
 IVKDEISSTAARIEKVWGENAMCYFISGCVTRQEGIGKIITDVRLSHMGPKNQYVYDNY
 RQLIHILIFRLLTGVS DNTSNILVGDGGNLFSVDENYVGAKDPRTALENRKIKELQLLL
 KTSFKVNKVTKEDIDSLPSWLFDTSKSDKIMNGVCNIGKNMGIGPTTLDIVKNNCTCIL
 GVVNDLLYDNK
 (SEQ ID NO: 9)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| Q9Y1H4 (Q9Y1H4) GAG | 43 | 0.011 |
| RAP1_YEAST (P11938) DNA-BINDING PROTEIN RAP1 (SBF-E) (REPRE | 41 | 0.031 |
| SLY1_YEAST (P22213) SLY1 PROTEIN | 40 | 0.070 |
| P90603 (P90603) MUC.CL-1 | 38 | 0.20 |
| O52224 (O52224) GLUCOSYLTRANSFERASE (EC 2.4.1.5) (DEXTRANSU | 38 | 0.35 |
| Q20202 (Q20202) F40E10.5 PROTEIN | 38 | 0.35 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 1092 of CT1039: this corresponds to nucleotides 133866 to 134957 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=1255

5'stop=1277

3'start=1353

3'stop=1375

5'primer=AATACATTGAAGCGAGAAGCTGC (residues 1255 to 1277 of SEQ ID NO: 8)

Tm5=58.71

3'primer=CCTTAATTTCCATCGGTTTCAGAA (residues 1353 to 1375 of SEQ ID NO: 8)

Tm3=58.18

probel=TAGTCATTGCGTAACGAG (residues 1278 to 1295 of SEQ ID NO: 8)

probelstart=1278

probelstop=1295

direction1=Reverse

Tm1=69.03

score1=1.96

length=121

CT630

Nucleotide

Genomic coordinates:

Start: 249215

Stop: 247358 (SEQ ID NO: 10)

Amino Acid

MALKDAFTERLVVNKVGSGTDMAPVVEDDRQKSLFQKVENLYRVLVVEQKNSAITLSGNK
 NTNKRQCRQVEEDKVFEGEDRTVSNLPQAVKETIAANAESILDYWKYKNIPLDTKKER
 SGKSDTFLRTAVICLVRCVSYKDMKTCSLIYEFCHKILNKSTLDPLLKDILDNKQELLH
 MDSKYGSKTTSPELAKETIEALYTTVYNHWTNAFKLYQASLTHKPVGTGKKYASVIHFIRT
 WRKIVKAYVSKHNNVERDLSLKNIMKNESADNANVLTIEKMYKKIGNSVKNTNNNSAHQM
 SDEDDDDDDDDDDCEGMDVDCDEASEREKKHQESLYPINTPVTTITGDYIFKVLELVLS
 HIHPEWKIPMCDFVNRNIPKLMKAMETDISNAVIEVRASKVNPVQILPIAANFWDCKSG
 KPPSDVKFCMMFNEPSSNETLSSGAGVFGRFIGGPFSHKSKELDIISNCLRSLLLNEAD
 NLSTRIWREGGSVVCFNYPITARGAVLGYGEQLSERSIKALWAKKIQDAVTESVKRQRN
 AADKNSRNCDLLGDEGVVSMKTVTFGCANMLKTQNGMGKFNVVVSFEDSIQANKEGAARO
 YMSQQVFTHSFPALDQKG
 (SEQ ID NO: 11)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| O97318 (O97318) PFC0865W PROTEIN | 45 | 0.002 |
| YB00_YEAST (P38114) PUTATIVE 126.9 KD TRANSCRIPTIONAL REGUL | 44 | 0.003 |
| Q08281 (Q08281) CHROMOSOME XV READING FRAME ORF YOL138C | 42 | 0.010 |
| Q92271 (Q92271) 12.8 KBP FRAGMENT OF THE LEFT ARM OF CHROMO | 42 | 0.010 |
| O62235 (O62235) F36F2.3 PROTEIN | 41 | 0.018 |
| Q83970 (Q83970) (CPV) | 41 | 0.023 |

CT631

Nucleotide

Genomic coordinates:

Start: 264975

Stop: 259164 (SEQ ID NO: 12)

Amino Acid

MGSKRPCSSGQEPVTKKQKNNNNNSNPVIVINIKSYPFLLATRTQVLRSAVAAAAASPSG
 SSSSSSSASAVKLPDTCKEARKVLSTVSLQQSLAVRYLCNSISVSYAGGGISVFHLGGL
 PGAGKTTMVKELIAVLNDHGLIDSGSADMLLCKKSNAKESLMCACKKPGGSSLMYPESV
 FSTLNKGFEIPVIFRKDEITLEKIQFVADKLKWKVIQVLNLRFLVIDEYTMASCRELVF
 IDAVLRIAKHRPDIFFGGVFVILLGDNRONSAVVEDNTNHIQKKIKNPSEEEKPKQNNKN
 NKNKKKKKEKKEKGEEEEEGDENEEEEEEEEEEEEESDDEAETKKEEEKSTFFQGSVEQD
 NFGQEDNAKLYTEVFILKLMFCSDFFGNPSNLRNIVNKRHEAILMKSNNVKSNNNLV
 SSAIKVEDCGNASNKKEVTAPSSSPAQSTAEENCDEFDDEEDDDFFNNEAFLKLMERNAL
 EKDRASGALNGFSLRCKSIDANEKIRSGTTSVSDKSSLDMMKSLPLSALIEEGICSEL
 AHISELKMSNANLEKYTENVCISIVFDMMAKAMREIDYSGREKLYIVSSLSERFKDTHLT
 SLMDDEILNVKYVHSGDPKCIDAVPFNSAHNRASAVAACVRNAFFRDGKDFVDETPIANY
 FKDNLRITVASFLENLTLYKELLAKSENIRSILLKKETGNNNSASSRTAAAAAYYEDDED
 YCYFDEEEAMDLEDGGSGSGSMKSSGGGGDDDEESGEMIRYTDIPDKLHRDASTLDRVG
 HLVDHFHVWKKWLTENKPSDLVRARVWYLYTLVRMQQVKFDNGKLPDLKLSALSGRLFHS
 PSEWATSTGVGVGGGGAGADKPLHDEYWLRLVLSMPISTGGDVGKSMLLPAYSSYLSALS
 RTYIMSSLRKRIIDIKHAYSLMYGISLFDMTANLQDLVDTRMAGRSSRNKSVFMDNFDPVQ
 YFDNIFPSMVNEFLMYRKEDVFNNQMMEGVKGSLKISRVLQTAHTENNNNNNRHNNSLK
 YSEKSIVLAMQMVTSISKGNERRKKIEEFITKEQGQPKDMCERLMANSKAKQEKDAISSK
 TDKMMGAITLTKKHVLKNAVSNLVDTSIIKETKKNNNNSSSSSTSLAAAAAVENSVPAL
 RVEVKFVVLNMDLSDISHEKTISHKYRQQLINAIKTRSTPLFDKFTDRKILRAAESPRAL
 TTILLDEKKKVTRAKSITLYQGQNVIFTTSNRMHGTQERFVTKDTGVVTNLMYKNGELT
 VFVYVERLGQKCLEIKEGRQIIGNPNIKNGGFGNNVYVQYLPFESSQAMTIYSCQGHTEFF
 RDTIVDLSGASTQDAYVAVTRNSNPQNLFIQNHVSVERGNLCNIKCAMSKDKAYTMPIGG
 IADFNQSDFINHDTVSVSREVAESSAAMDDDYNGDGGVTMYSAYDPSKDVVAAAEFFILS
 RSGKSLSFNASWMANTAKVIOQHGLETELKNIRDFFFGVNNGDVAKHYEKLCKNKKMIELY
 TAIVRSITHYSIASGIVKQPSSKLCEEYETKQKNKKDYIKIHPVFVNRAPESTIEMLLF
 DIAPHNKATIVQFYVHYIFLVYEKLVNLSSFAFLPSNPCLNQYVRPKSITTNSTHVP
 NLGYESKDFAHCKDGGGERDVKLRLPITSADEFSNNIEGILKKVSDTSNQNKVNKYMDVVC
 KSMQHNLRRTGKFCRPTETCGLSKHGSIVTSTCTAQEKGENIHVDAEKGLCMSDEANVY
 CMLMFMSKIAAASGVSEFFPIKDKSILESNPETPSDTISLLAPRKTIPTNNLHFSMSDEV
 LECCQVHPMKRVQFSLHVKRTGGALKSTFEEEGLPTKIFSPNFATYPLFKKCKMYGAI
 IAMTEMQGHFAKYSTLDIRKSMFTGVGTVDLEKISGEGNEVMDKVDKFIKVNVSNIIF
 KEQGKRVSFFVSCAIH
 (SEQ ID NO: 13)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|--|-----------------|------------|
| Q9YTL7 (Q9YTL7) ORF 48 | 68 | 5e-10 |
| O35788 (O35788) CYCLIC NUCLEOTIDE-GATED CHANNEL BETA SUBUNIT | 68 | 5e-10 |
| NAB3_YEAST (P38996) NUCLEAR POLYADENYLATED RNA-BINDING PROTEIN | 68 | 8e-10 |
| Q07034 (Q07034) RNA BINDING PROTEIN | 68 | 8e-10 |
| O96229 (O96229) HYPOTHETICAL 78.6 KD PROTEIN | 66 | 2e-09 |
| O96134 (O96134) SER/THR PROTEIN KINASE | 66 | 3e-09 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has sequence identity to nucleotides 1 to 830 of CT631: this corresponds to nucleotides 259304 to 260133 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=2408

5'stop=2429

3'start=2489
3'stop=2509
5'primer=GAGCCCGTGTGTGGTATTTGTA (residues 2408 to 2429 of SEQ ID NO: 12)
Tm5=58.49
3'primer=AGCGTCCAGACAGTGCAGATT (residues 2489 to 2509 of SEQ ID NO: 12)
Tm3=58.56
probe1=CACTAGTGAGAATGCAAC (residues 2432 to 2449 of SEQ ID NO: 12)
probe1start=2432
probe1stop=2449
direction1=Forward
Tm1=68.99
score1=1.99
length=102

CT632
 Nucleotide
 Genomic coordinates:
 Start: 268485
 Stop: 267717 (SEQ ID NO: 14)

Amino Acid
 MAGVDLYGGHIKPYGETVFNNKMQGNRGKIRALINEKAAATLPMSEDNISAWVTEVAADV
 FPDPKSALTFFVPNKSLNAFAWDVLKTPASVEIDIGKRIPQLIENLHMSDFTVAIFRVKC
 DDQGRYETSYNLSPSMGGKINHGLIRTLAKAQDIVVWKRDFSLTIENFEVDNGKKRLDFL
 FNNQTDKSCFVKIFHEMESEKDIAIKPEKRGSSAVWDEVYSDIVTKNTRNAKFSRLRYNE
 KPVDHLLLYCMVTYF
 (SEQ ID NO: 15)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| 097036 (097036) PLC-BETA2 | 34 | 1.5 |
| 045329 (045329) F09C6.2 PROTEIN | 33 | 2.5 |
| P97868 (P97868) PROLIFERATION POTENTIAL-RELATED PROTEIN | 32 | 5.6 |
| P70287 (P70287) RETINOBLASTOMA BINDING PROTEIN 6 (PACT) (FR | 32 | 5.6 |
| AAD49229 (AAD49229) EHEC FACTOR FOR ADHERENCE | 31 | 7.4 |
| CAB55629 (CAB55629) LYMPHOSTATIN | 31 | 7.4 |

Comments:

TaqMan Primer/Probe Sets:

5'start=378
 5'stop=403
 3'start=452
 3'stop=477
 5'primer=TGAAACCAGCTACAATTTATCTCCTT (residues 378 to 403 of SEQ ID NO: 14)
 Tm5=57.68
 3'primer=CCTCTTCCAGACTACAATATCTTGGG (residues 452 to 477 of SEQ ID NO: 14)
 Tm3=59.43
 probe1=TCAGAACTGGCTAAGG (residues 434 to 451 of SEQ ID NO: 14)
 probe1start=434
 probe1stop=451
 direction1=Forward
 Tm1=69.15
 score1=1.84
 length=100

CT633
Nucleotide
Genomic coordinates:
Start: 266790
Stop: 266442 (SEQ ID NO: 16)

Amino Acid
MASPLVASLGGGKNILFGLLLITIIIVIVIAVIIIKAPLLASLLAGTALAGTIASALGSIP
GVGGAFKKAFGKGGKGGPKTPDGGAKKTNQKPKKGKKKPTRRSIFKRIPKIKF
(SEQ ID NO: 17)

Top Blast Hits

| Sequences producing significant alignments: | | | Score (bits) | E Value |
|---|----------|---|-----------------|------------|
| Q02391 | (Q02391) | CYSTEINE-RICH FIBROBLAST GROWTH FACTOR RECE | 36 | 0.092 |
| Q91019 | (Q91019) | MUTANT CYSTEINE-RICH FGF RECEPTOR | 36 | 0.092 |
| Q69526 | (Q69526) | GLYCOPROTEIN B | 36 | 0.12 |
| Q9XZ15 | (Q9XZ15) | HYPOTHETICAL 29.3 KD PROTEIN | 36 | 0.12 |
| Q14113 | (Q14113) | AORTIC CARBOXYPEPTIDASE-LIKE PROTEIN ACLP (| 35 | 0.21 |
| P79922 | (P79922) | MODIFICATION METHYLASE (EC 2.1.1.73) (CYTOS | 35 | 0.21 |

CT634
 Nucleotide
 Genomic coordinates:
 Start: 283360
 Stop: 282673 (SEQ ID NO: 18)

Amino Acid
 MVSSRTSTTSSAVAATSTLLPTKRKREPEEVKVKVEVKMEQEELVEDSSSNKRPRIKEE
 KEEHHKETHHLSLPCKEEEDDGESEEEYEEEDDEEYEDRVDDDTAEKMENLLVQLDNTTK
 NIKLKNPLREHDMVSHYEHEFEVQNTVNFSGVLSDIGFLINREAVSRWGNTPPPKEFG
 DMEIGSLTVNQLLHKCDNFVQAVVQVKEDITPSIEVTIDSLIDDPW
 (SEQ ID NO: 19)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| O35788 (O35788) CYCLIC NUCLEOTIDE-GATED CHANNEL BETA SUBUNI | 60 | 2e-08 |
| GARP_PLAFF (P13816) GLUTAMIC ACID-RICH PROTEIN PRECURSOR | 57 | 8e-08 |
| Q9YTL7 (Q9YTL7) ORF 48 | 56 | 2e-07 |
| O94922 (O94922) KIAA0835 PROTEIN | 55 | 3e-07 |
| O08995 (O08995) MYELIN TRANSCRIPTION FACTOR 1 | 55 | 3e-07 |
| Q93424 (Q93424) HYPOTHETICAL GLYCINE-RICH 37.0 KD PROTEIN E | 55 | 5e-07 |

Comments:

EST confirmation of the predicted transcript:
 An isolated EST has equence identity to nucleotides 575 to 1 of CT634: this
 corresponds to nucleotides 282678 to 283252 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=210
 5'stop=231
 3'start=287
 3'stop=308
 5'primer=CCTCTCCCTCCCATGTAAAGAA (residues 210 to 231 of SEQ ID NO: 18)
 Tm5=58.19
 3'primer=TCGTCGTCCTCTGTCTTCAT (residues 287 to 308 of SEQ ID NO: 18)
 Tm3=58.40
 probel=AAGAGGAGGATGAGGAAG (residues 266 to 283 of SEQ ID NO: 18)
 probelstart=266
 probelstop=283
 direction1=Reverse
 Tm1=68.99
 score1=1.98
 length=99

CT635
 Nucleotide
 Genomic coordinates:
 Start: 285773
 Stop: 284075 (SEQ ID NO: 20)

Amino Acid
 MARSVGLLSVTPEYDTFKYIKMEEFKTLKVKNNGFTISGENPKYEHILLSFKSVDRVTKS
 ELRDGLYIVRLKDKEVLHIKNGVHRLRQLTGDNTLQVGLKYTHNLPRLGSLQDDGCEDY
 GEKWNESLPIDMQNINKIVKEKALLSDKNFKFSPLYRLLHERLSNAAVKKCDYMIITDF
 LVGCGYTPSHCPRTLNRMEQLLVEQCGFSSRISVYDICDRLTYKGAYIANPITGSYSNMC
 LIVPMDKLGLIFYNSTHPSAKSIGNYMSSLFNATVIYANERDNLQMDNFRREIKFAENEV
 NMKEEELKELRKRCVSEEQRISLRDVHKKSSIATSRYGACLVFAFSRDRFSLLCRTN
 GNGSFYSATEEGIRYVSSPEYKKRDVGERRPRLIMSI TGS DAPICIRDSVRNHFKTRLFS
 RTSGNSITFAVPPGERELMEMVREVTGTDIKIFMDNGKVYQNGAEINVIDPTSKEYKELL
 KREENLPEDERKRLRRERRMIFNTSRAISMYNEERG DGGSGGETS EDGDGNGSTSSKGEK
 RKREENEGNEYVVLNKCACKDIKVC
 (SEQ ID NO: 21)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q21885 (Q21885) COSMID R09H3 | 37 | 0.32 |
| O77336 (O77336) PFC0425W PROTEIN | 37 | 0.42 |
| O00164 (O00164) RIBOSOMAL RNA UPSTREAM BINDING TRANSCRIPTIO | 36 | 0.71 |
| UBF1_HUMAN (P17480) NUCLEOLAR TRANSCRIPTION FACTOR 1 (UPSTR | 36 | 0.71 |
| YPT2_CAEEL (P41880) HYPOTHETICAL 21.6 KD PROTEIN F37A4.2 IN | 36 | 0.93 |
| YG2K_YEAST (P53253) HYPOTHETICAL 106.5 KD PROTEIN IN CTT1-P | 35 | 1.2 |

Comments:

EST confirmation of the predicted transcript:
 An isolated EST has equence identity to nucleotides 501 to 1 of CT635: this
 corresponds to nucleotides 284047 to 284547 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=686
 5'stop=706
 3'start=799
 3'stop=822
 5'primer=CAAACCCAATCACAGGCAGTT (residues 686 to 706 of SEQ ID NO: 20)
 Tm5=58.55
 3'primer=GGTTGCATTGAAAAGAGATGACAT (residues 799 to 822 of SEQ ID NO: 20)
 Tm3=58.19
 probel=ACTCCAACATGTGCCTAA (residues 707 to 724 of SEQ ID NO: 20)
 probelstart=707
 probelstop=724
 direction1=Forward
 Tm1=68.89
 score1=1.89
 length=137

CT637
 Nucleotide
 Genomic coordinates:
 Start: 300432
 Stop: 299085 (SEQ ID NO: 24)

Amino Acid
 MGGEDSFDDRYDS DALWENEGAKSIQVKETDLEVYRMHRAVPTLEEKNTALRYSDWS
 PVYRVPLFSLKDGSDPHERDFSNVDPRREGKVPVKVRRVDVRNPSRTAAIFVPTGPGHLH
 VSSYTG DGM LVCPNHNFIGDLCSEIASDITIYNTSSSGRLSYATNFNSVEDNSPVGILFE
 TLPDDKMFOQVSIFSATEPASNISIGPMSHVKIKLGYDEENATAVGVIRYGGGLFYTSVG
 ACIIPEGVFFDDVVG NHSSMNIYNMTNPKEIVLKEPRGEDAMEEDDGEEADYNFLGYVV
 RFEHDLKMQAMSSAYSSVSIDINSSSFHKCFLIKPKYNSILQPLVSSEVV LNDLSLNTRG
 REVEFHDRLPSGAQDNSYSIVKYMKIVSLKEGLKVVNPIINTELYKKKQALKVHVLNMTR
 DVVGLDTSEHSFGVIVCHAAKLPEVIGQ
 (SEQ ID NO: 25)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| O23592 (O23592) CARBOXYL-TERMINAL PROTEINASE HOMOLOG | 34 | 1.6 |
| YGA1_YEAST (P53199) PUTATIVE 3 BETA-HYDROXYSTEROID DEHYDROG | 34 | 2.1 |
| VILI_DICDI (P36418) PROTOVILLIN (100 KD ACTIN-BINDING PROTE | 33 | 4.8 |
| Q9Z5B9 (Q9Z5B9) PUTATIVE TRANSFERASE | 32 | 6.3 |
| Q9YUY4 (Q9YUY4) ENVELOPE GLYCOPROTEIN | 32 | 8.2 |

Comments:

EST confirmation of the predicted transcript:
 An isolated EST has equence identity to nucleotides 3 to 1254 of CT637: this
 corresponds to nucleotides 299132 to 300383 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=678
 5'stop=702
 3'start=788
 3'stop=812
 5'primer=CGGTGTCATTAGATATGGAGGATTA (residues 678 to 702 of SEQ ID NO: 24)
 Tm5=57.26
 3'primer=TCTTTTGTTTGTTAGTCATGTTGT (residues 788 to 812 of SEQ ID NO: 24)
 Tm3=57.75
 probel=CTACACCTCTGTCGGTGC (residues 705 to 722 of SEQ ID NO: 24)
 probelstart=705
 probelstop=722
 direction1=Forward
 Tm1=68.85
 score1=1.85
 length=135

CT1040
 Nucleotide
 Genomic coordinates:
 Start: 137588
 Stop: 139940 (SEQ ID NO: 26)

Amino Acid

MAAAAVSGEGRISADLLLLLEQLTPDGDVIRYDSEQYTKPRKIFGDKSVIETIGHFLIHH
 NQGESYQIASSVLEKFPALLNCIWNGESGGMALWKALYRAKKYRLLSLLVHKIKNWPSV
 AVIPIYGSVCDREERPIIMSEIIDKETLQTICKSDIRSLGMMNAKHGTLGGNFLHFIYAR
 STKPFENFQYEAMGANAVLMAAEAIYDGRDHGLNPSEYTFPGLESADVYGNPVEIAIS
 GDDDNMLLNLCNYGVSYEKTRGRVNRSLDLFLKMNTASKCLSVLKFEKHFKIESNTPK
 GEFEKEAETCVNCLDRNNVLTKGSEQESYKLSGCHFLHVKCLRNICIVSQHLRCEKCLKR
 FDESILRKCTPNLNWWTMPAGAGNEEEICFMRNKKLVDDFRKLLSPVSIPIHFFKNSRQR
 NLDMLCPYSDHITIIPNKEDPKKNEDGNRVRVNHTAISEKQNKKEEDARIKRVAVRTFTAI
 REKONKEEDARIKRAVDMAVAANINEKNKEEDARIKRAVDMAVAANINNKEEDARIK
 RAVDMAVAANINEKNKEEDARIKRAVDMAVAANINNKEEDARIKRAVDMAVAANINN
 KEEEDARIKRAVDMAVAATNEKNKKEEDARIKRIIDLTVDMRIQRIVDMAIAAATKKDKK
 EEEKRTKREQLRADLRRAMDVMNEVQKKLEDMELEKGCNKDEAKNTSNVSSSSVVAYS
 KEIVPCLGNNNNAVIGMTSTNYSANNTKNNVFGSPHKFSFNDASRFSNIVETPKMSFNFS
 FKT
 (SEQ ID NO: 27)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q9ZU69 (Q9ZU69) PUTATIVE VICILIN STORAGE PROTEIN (GLOBULIN- | 71 | 4e-11 |
| Q26775 (Q26775) TB-292 MEMBRANE ASSOCIATED PROTEIN | 70 | 5e-11 |
| INCE_CHICK (P53352) INNER CENTROMERE PROTEIN (INCENP) | 67 | 5e-10 |
| YDF3_SCHPO (Q10475) PROBABLE EUKARYOTIC INITIATION FACTOR C | 66 | 7e-10 |
| Q26774 (Q26774) TB-291 MEMBRANE ASSOCIATED PROTEIN | 66 | 9e-10 |
| P91257 (P91257) SIMILAR TO C. ELEGANS UNC-89 | 65 | 2e-09 |

Comments:

EST confirmation of the predicted transcript:
 An isolated EST has equence identity to nucleotides 853 to 415 of CT1040: this
 corresponds to nucleotides 138912 to 139349 of the genomic reference sequence.

CT1041
Nucleotide
Genomic coordinates:
Start: 140110
Stop: 141616 (SEQ ID NO: 28)

Amino Acid
MVYKGFVSPSFITIRMTSNRPTTSPLSFSEGFSLSGDKYDTYEDILLEQFNCFKTSSPS
SARKSEIEDKTLIFQLKEGEKFHLAKGIEELREILDDNSATIEPIISPTTFNDRNELLNH
EGDISSSPLYTQIMKHISPEHDIYELDLIVGTDLLFGLGVNLRNVSKLMKKISYGTINVV
DVCHRKFFNNRIIVNPISSEFSKNVCIIPLFSAEEFSSLGECRDLFNGICDDVERYINS
YFFYPENTTTTTTAPSSPEMEIADEEEQSPKTIKRNDNASRNWSGVCLIFEVFKNTYYI
INRGDRGGSFEKAVKSAISSIKEKRCKITDINGNKPRLVMVITGCTELYFKDALKQIGE
NRRKFLKMNGNYFSLIDEQADLIEFAMSVSGAGERIFVNGLGMFQNRKMIPIVIDPLTYEN
VVCGEHDIQKEDAILSVRRAIADYNDFVSKNKRKGKRSAAEENEDEDADASSSSSSSPPP
SSPPAHKKSRLPDEGEKCTL
(SEQ ID NO: 29)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q26258 (Q26258) BR2.2=BALBIANI RING (5' REGION, REPEAT UNIT | 38 | 0.12 |
| Q23804 (Q23804) SPID PRECURSOR (FRAGMENT) | 36 | 0.82 |
| O97324 (O97324) MAL3P8.1 PROTEIN | 36 | 0.82 |
| Q99112 (Q99112) HOMEODOMAIN PROTEIN BW2 | 35 | 1.1 |
| Q99111 (Q99111) HOMEODOMAIN PROTEIN BW2 (FRAGMENT) | 35 | 1.1 |
| YA55_METJA (Q58455) HYPOTHETICAL PROTEIN MJ1055 | 35 | 1.1 |

Comments:

TaqMan Primer/Probe Sets:
5'start=822
5'stop=846
3'start=929
3'stop=950
5'primer=AAAGAGAAATGACAACGCAAGTAGA (residues 822 to 846 of SEQ ID NO: 28)
Tm5=57.46
3'primer=GCACTCTTCACAGCCTTTTCAA (residues 929 to 950 of SEQ ID NO: 28)
Tm3=58.24
probel=AACTGGTCTGGTGTCTGT (residues 847 to 864 of SEQ ID NO: 28)
probelstart=847
probelstop=864
direction1=Forward
Tm1=69.07
score1=1.92
length=129

CT1042
 Nucleotide
 Genomic coordinates:
 Start: 141695
 Stop: 142541(SEQ ID NO: 30)

Amino Acid
 MAVNLDNVLVNINNKDEDLTKLVSEAIKRRAKTVFDTKNQAGFDMRRQVEAALYEAISKKK
 EKAIAFDELIQERGDEITPLTTMQYEEWVNRTITPSLTENLLGDVEHADFLDRMTPVS
 EEDIEGFAASTFKEVSDSKTATVIVKADCETGDIDEVYNLAPSGVTQEIKIYRSNNSSEL
 DNVADSFHIYKISATDSDSGNTKKLLYGLRNKKAGYTCLCRIFAEIESDGIMANTNIGVAE
 NNRDEIDENEEGKYGFLIPKQPAGAKLIYFFLNCWTX
 (SEQ ID NO: 31)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|--|-----------------|------------|
| gi 2492980 sp Q10043 YRP1_CAEEL HYPOTHETICAL 37.6 KD PROTEIN R1... | 33 | 0.79 |
| gi 401691 sp Q00933 YSCI_YERPS YOP PROTEINS TRANSLOCATION PROTE... | 32 | 1.8 |
| gi 1706480 sp P51892 DNL1_XENLA DNA LIGASE I (POLYDEOXYRIBONUCL... | 32 | 1.8 |
| gi 267570 sp Q01250 YSCI_YEREN YOP PROTEINS TRANSLOCATION PROTE... | 31 | 2.3 |
| gi 3121979 sp O07597 DAAA_BACSU D-ALANINE AMINOTRANSFERASE (D-A... | 31 | 2.3 |
| gi 6686325 sp P71018 PLSX_BACSU FATTY ACID/PHOSPHOLIPID SYNTHES... | 31 | 4.0 |

Comments:
 EST confirmation of the predicted transcript:
 An isolated EST has equence identity to nucleotides 700 to 1 of CT1042: this
 corresponds to nucleotides 141748 to 142447 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=350
 5'stop=367
 3'start=392
 3'stop=414
 5'primer=GGACCGAATGACACCCGT (residues 350 to 367 of SEQ ID NO: 30)
 Tm5=57.70
 3'primer=CCTCCTTAAAAGTAGAAGCAGCG (residues 392 to 414 of SEQ ID NO: 30)
 Tm3=57.66
 probel=AAGCGAGGAAGATATTGA (residues 368 to 385 of SEQ ID NO: 30)
 probelstart=368
 probelstop=385
 direction1=Forward
 Tm1=69.01
 score1=1.88
 length=65

CT1043
Nucleotide
Genomic coordinates:
Start: 142610
Stop: 143699 (SEQ ID NO: 32)

Amino Acid
MTVLAVYTAPQIKKSKKRKIEDENEEEPVKLTLEDFVKGRLLNAVKEKPAEYFELLISADT
EAALKTAETALRDFVIENDSVEIDVEEVLEEKPREYVFKLAGATSETLTNTIIAEVQKK
AALITEEDITIKMLKQFRAANKDNKDGEATPEEKEDFTNNSDLVGLYLNEVEKTTNIVI
NKIFPHEMVFERCAILIEDFTDGVVTDQAIQIPSNKYKIRLVEGDEPEVFPGLDCLDLAVS
VDKINHVLKISAKNGCENNCFVIIIPRFPVPGSVSSMILGSTDQVKPKTFLFLANKNDSTH
FQFTMDKQHSVGCELDMLIFSERNLRLNPDSKPRPLSDADILASYGKRLGTGVFTTENLV
DD
(SEQ ID NO: 33)

Top Blast Hits

Sequences producing significant alignments:

| Score (bits) | E Value |
|-----------------|------------|
|-----------------|------------|

| | | | |
|---------------------|--------------------------------------|----|-------|
| Q26938 (Q26938) | KINETOPLAST-ASSOCIATED PROTEIN (KAP) | 40 | 0.022 |
| O23332 (O23332) | CENTROMERE PROTEIN HOMOLOG | 39 | 0.037 |
| Q9ZES5 (Q9ZES5) | CTC PROTEIN | 39 | 0.049 |
| AAD45753 (AAD45753) | ANON1A3 (FRAGMENT) | 38 | 0.083 |
| O66878 (O66878) | CHROMOSOME ASSEMBLY PROTEIN HOMOLOG | 38 | 0.11 |
| AAD45759 (AAD45759) | ANON1A3 (FRAGMENT) | 38 | 0.11 |

Comments:

EST confirmation of the predicted transcript:
An isolated EST has equence identity to nucleotides 1 to 982 of CT1043: this corresponds to nucleotides 142640 to 143621 of the genomic reference sequence.

TagMan Primer/Probe Sets:

```

5'start=543
5'stop=567
3'start=635
3'stop=659
5'primer=CAAAATATTCCTCATGAGATGGTT (residues 543 to 567 of SEQ ID NO: 32)
Tm5=58.37
3'primer=CTGATTTTGTATTTGTTGGAGGGTA (residues 635 to 659 of SEQ ID NO: 32)
Tm3=57.33
probe1=GACTGATCAAGCCATTCA (residues 615 to 632 of SEQ ID NO: 32)
probe1start=615
probe1stop=632
direction1=Reverse
Tm1=69.01
score1=1.98
length=117

```

CT1044
 Nucleotide
 Genomic coordinates:
 Start: 143759
 Stop: 144689 (SEQ ID NO: 34)

Amino Acid
 MSSSSSETPKTSTDTGEERIKDIVNALDNNGEWLSSYIDPIINNHISRKTAETVQKINQE
 VDERYDRKIADKINEIKSSIFSAQTMVDQYAITDFQEGKGANGTGPMGPVNTVIDTTL
 NKMRGNMLEYAEDMWDGDDWKRFSSSMTTLEFDLSYDLTMMRGSDGYFAFPFRGTTKKIK
 MDGSRKKEDPINCIIISVTYPNKVGDWEWEEKEREVNFNLERVDDYERDIHVSILCMLHAQ
 LDNFEQALGENANSFYFKKGQRMFLPKKSKLFNRPTVEDSDMFIIFFPPASDQDFADDI
 YYRIIVTCS
 (SEQ ID NO: 35)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| O13706 (O13706) HYPOTHETICAL 11.8 KD PROTEIN C13F5.07C IN C | 36 | 0.47 |
| Q20960 (Q20960) COSMID F58A6 | 35 | 0.61 |
| CAB59514 (CAB59514) HEAT SHOCK PROTEIN 70 | 34 | 1.8 |
| Q45851 (Q45851) NEUROTOXIN TYPE F | 34 | 1.8 |
| O17208 (O17208) C01B12.2 PROTEIN | 33 | 3.1 |
| BAA83026 (BAA83026) KIAA1074 PROTEIN | 32 | 4.1 |

Comments:

EST confirmation of the predicted transcript:
 An isolated EST has equence identity to nucleotides 1 to 954 of CT1044: this
 corresponds to nucleotides 143768 to 144721 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=330
 5'stop=350
 3'start=406
 3'stop=427
 5'primer=GCCAGTGAACACGGTTATCGA (residues 330 to 350 of SEQ ID NO: 34)
 Tm5=59.57
 3'primer=ATCGTTTCCAGTCATCTCCGTC (residues 406 to 427 of SEQ ID NO: 34)
 Tm3=59.04
 probel=CTCGAATACGCTGAAGAT (residues 382 to 399 of SEQ ID NO: 34)
 probelstart=382
 probelstop=399
 direction1=Forward
 Tm1=69.00
 score1=1.99
 length=98

CT1045
 Nucleotide
 Genomic coordinates:
 Start: 150687
 Stop: 154344 (SEQ ID NO: 36)

Amino Acid

METTMDNVVQNNVDVTKPTPDVATVTTATEKRQSCKEKKDQLKAECPOVLRAKLKLSNTLKA
 NFGKSMSAIFAQHLVDMTNAKHFKDPKTKKILELDGSSSSDSEEEETSSSSSKRKRGS
 RSASSKKEKCPNTIKNWLNDAAQGVFRQFADIIINLPSFDDLREVKDEQTELKTIYDLYR
 QDMEKVVVEVLGRQDLFDHKSEIAKGLARFDTHVSLLPDSORSAVLDSISKELEKNSKGP
 NSNIFDTLNTLKEEIKELLCHHVKYLLQNLTPEANFVFNSSVKYVKKSYYIQTSEME
 SDEFKSLLTGVNIKILEKIISSDNNVATPYKHITNPRNIISLQKVRETKPVSKDYPFRV
 DTARDIVLLPETGGISDLPIKPVTLQLVSYINALFSLERRNVFTDGGFFNAACVLISQCL
 TNANLLSNDFFPKPIELAAVTRHNLLSMKMLQEGSSSEKSKKKEKKKKKKGGGGGDDDS
 DSETDSSSSSSSSSSSSSSSEDEEEKGEAVEKGGKTKRKTKKKPSKDDDLDTISKLI
 LKTGGYFHDTSSELGNKIRNLIKDDDFAGVAQYAVTITEMQSTPMNQRLVSSLLDLIMRLK
 EQVKYSVDTESTSTAKSNNALDSAKLTSQQVVTMMVDSGAELARLAFFVVDNTVNVN
 RHEAFILTSKLLPSNENRGLKTVVESEFFKNLTISNKVSTNEEMMSVMPFEDEQQQQQCP
 QHEQQPDLKRVVGEVFLEMGKSI VNSFPNSKSVQLTADAFKQNYSPMGRINLAAKIKTA
 ISIGSNISPNILFSNLPESVGNNTVTGLRLTNLLKNISQSAQANNIKNANTLVNNTMDQ
 QNSAAMSILLFPPTSKETSIFPGNDPSSIKLQDMTMSNLARGFYSAEGCIGVVRREF
 DEGGVKAYTLLVDSNTMDMAVNFAAQSLKSMSEALTNNANMNPSNVLEGGSFVDGALS
 YMEKNGSDCEPTPLAKYTMKDVSRYLKKFNNDKNTQDLYKNRAERALVEQVTNKPTSVV
 HSQLANAMGVAVIGAASIKLMEAAEAESEMRAANYQATSKSTNAINITNTIGMIRNTTHL
 CTTIAVSAADMSKLANNHFMVSLNTANNSHSSRRGDRSSLMLQQQPTHSFLEQTRGR
 GGGVLGSGTEQTKDHVERMKRDWILNMISPEDKNTTTTTTPSNAGRTLGYGSNITGINTIK
 QDDKSMMDKLSEMSFRT
 (SEQ ID NO: 37)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| Q9XYZ6 (Q9XYZ6) HYPOTHETICAL 75.5 KD PROTEIN | 59 | 2e-07 |
| O55035 (O55035) PEPTIDYLPROLYL ISOMERASE MATRIN CYP (EC 5.2 | 59 | 2e-07 |
| VTA2_XENLA (P18709) VITELLOGENIN A2 PRECURSOR (VTG A2) [CON | 57 | 5e-07 |
| Q07034 (Q07034) RNA BINDING PROTEIN | 57 | 9e-07 |
| NAB3_YEAST (P38996) NUCLEAR POLYADENYLATED RNA-BINDING PROT | 57 | 9e-07 |
| O95367 (O95367) CBF1 INTERACTING COREPRESSOR CIR | 57 | 9e-07 |

Comments:

TaqMan Primer/Probe Sets:

5'start=1624

5'stop=1646

3'start=1702

3'stop=1720

5'primer=AAAACAGGAGGTTACTTCCACGA (residues 1624 to 1646 of SEQ ID NO: 36)

Tm5=57.88

3'primer=CTGCATATTGGGCTACGCC (residues 1702 to 1720 of SEQ ID NO: 36)

Tm3=57.81

probel=CACGAGTGAAGTCGGCAA (residues 1647 to 1664 of SEQ ID NO: 36)

probelstart=1647

probelstop=1664

direction1=Forward

Tm1=69.00

score1=1.99

length=97

CT1046
 Nucleotide
 Genomic coordinates:
 Start: 154556
 Stop: 156932 (SEQ ID NO: 38)

Amino Acid
 MSLVENNTQEEMILETTVEGVVEGAEVAPRGVVRPLPSSSSSSASDSEDDEGGEQPQTK
 PPKKKRNINSGKYWKIETIEPASPEMLSAVNDIDNVSKTIPLIDNSFGVQFKKSVSEEQI
 KTLTETIATIVEYGTITNVKYSTFNQLERTGEPLKKRSNNGNNYRYWQIRIEAAAAENV
 TQAVLDAIVEGNDTVIKAILLPEGEGIGLQFNKSVSSQAKNIVQAADIEFGQVAHMKCN
 LFHKMEKADESSNSSGESPKVKVRRNKSQPTNSYYTFTMIGDSLQERIDNAIKVIEMSP
 VKRPFNSAAAAEEDTTTTTTSTGVVNPRGIKDIHFFDSSISKGCFTVRNIVAANGEVP
 QEEFVSELYTNLLKVEEKVDHPTFKKLIHDRTMNRHIKAWYCICPYTTGGVPPAADKVS
 AKGIATYRIYEDRTGVFQFDGAHTSTTPAQAAEATGAIHKSMLFQSPGTDIQKFLDAKKA
 EGLEPISSGEIVYRSKWSPNDSRATRCFKFYSSSDEKMNIADVLSIVHTDGLFSSVHFRK
 DTMEYGVAKSKSKIIPKTIKIKKGGDTFHSEEDIEVPVKFTAITSEELNRECNTKGMNSL
 RAHKRRKSNSTTTTSTSTSTTANTPKKTKKSASASDPFAKLTLDYVDSTSFVFNIS
 KEMVQRILAQERVTKLAVKNEEKMEIVEGEEAQETYRGIVIKITNAKAYNLANKTSGVL
 FPADKVLKHTLEDLGDVLDVDFVREDNVNKTVASTTTTSSSENKASGGDDEETPMEFETD
 GEKLLHELLNE
 (SEQ ID NO: 39)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|-----------------------------------|-----------------|------------|
| Q45759 (Q45759) CRYSTAL PROTEIN | 41 | 0.023 |
| O97003 (O97003) L1156.7 PROTEIN | 41 | 0.030 |
| O02061 (O02061) B0041.7 PROTEIN | 41 | 0.039 |
| AAD55361 (AAD55361) XNP-1 | 41 | 0.039 |
| KI67_HUMAN (P46013) ANTIGEN KI-67 | 40 | 0.051 |
| P79065 (P79065) NOC1 PROTEIN | 40 | 0.066 |

Comments:

EST confirmation of the predicted transcript:
 An isolated EST has equence identity to nucleotides 1 to 935 of CT1046: this
 corresponds to nucleotides 155963 to 156897 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=1338
 5'stop=1355
 3'start=1419
 3'stop=1440
 5'primer=TACCCCAGCACAGGCAGC (residues 1338 to 1355 of SEQ ID NO: 38)
 Tm5=59.36
 3'primer=TGCCTTCTTAGCATCGAGGAAC (residues 1419 to 1440 of SEQ ID NO: 38)
 Tm3=59.46
 probel=GAGGCAACTGGTGCTATT (residues 1357 to 1374 of SEQ ID NO: 38)
 probelstart=1357
 probelstop=1374
 direction1=Reverse
 Tm1=69.01
 score1=1.98
 length=103

CT1047
 Nucleotide
 Genomic coordinates:
 Start: 159378
 Stop: 161256(SEQ ID NO: 40)

Amino Acid
 MSCSSSSSSSSSEENEVGVGGGGRIGPTEAKKKILRKRKSSVKSTSSSSSSSSSSD
 DSDSDREEKEGRKLYVDIADTRKPPKVRKLDTPSQTLENDLYMSSSSSSSSSSDSSSSS
 GEEESDDDDDDDDYDPDNVHVLGCKKEKSPQDIEAEKEKEEEYEEEFKRMALPSRINTSV
 DCVIPDRILTLFSTLLKNSFQFSQPVSFRLRLVMKQVNEAMNSAFSSMLSSSGMRLVEDS
 LGDTSKISSFITPQTDTSNSSSSSTFVNNCTDEDIKKRNIAMGRVAELLSNIAASSNEEN
 NFRPVVSLMRGPTCGGSNASNKKLNSNRQTIPQVLNKVIFFREIHSVIALYLSSVCVQRA
 MNNDNTNSSGYAEGMVTKILNIIGKIPYNEMSREKFISVGRDALYLYQNVTDMTGPKHN
 KRLRIPQQQADFCYIIAMLVNDVPITSDLLLTGKATNLVQFASAMVDPAYRLAVHKMASV
 FNSSYSVYKVLDDLHKMLLRANLILSILSARNKCLSERKPRTLTSVYLFLNHLRNKLR
 SSGLTSEESSLGTAVKLVSQQLMYEGVTRQTIEDGCSMISGNFEDEDGVTCLKCLGADV
 VKTVGLSALLSDRLRKNIRRVPFY
 (SEQ ID NO: 41)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| SR40_YEAST (P32583) SUPPRESSOR PROTEIN SRP40 | 74 | 2e-12 |
| NSR1_YEAST (P27476) NUCLEAR LOCALIZATION SEQUENCE BINDING P | 63 | 6e-09 |
| VIT2_CHICK (P02845) VITELLOGENIN II PRECURSOR (MAJOR VITELL | 61 | 3e-08 |
| O95815 (O95815) DENTIN PHOSPHORYN (FRAGMENT) | 60 | 4e-08 |
| VIT_ICHUN (Q91062) VITELLOGENIN PRECURSOR (VTG) [CONTAINS: | 59 | 1e-07 |
| GAR2_SCHPO (P41891) GAR2 PROTEIN | 58 | 2e-07 |

Comments:

EST confirmation of the predicted transcript:
 An isolated EST has equence identity to nucleotides 2 to 888 of CT1047: this
 corresponds to nucleotides 160277 to 161163 of the genomic reference sequence.

CT1048
Nucleotide
Genomic coordinates:
Start: 161717
Stop: 165020 (SEQ ID NO: 42)

Amino Acid

MFGSSANNFNGDKKSSSSSSAAASSDDQQLGPLGLSTADFKKVAAILANRTESLYLLPDS
PNFKNVINPNQISIVPFLGSSKAAESGSANKNENQAENSSKGGSDGKKSSQONKFNLLN
KVEAEEMAFKRVAELIADTPPSKDNPLRDDPDAIPSRNPWVKLTQKNLEYLFW EAVTIEV
SNDRSIRSGRYLQASEVGENPFLMTISVDIRILQRMALNVVWFFNRFFRMVSGLGVENRA
NSTYVATSDAIAQIWVEMLLKNFISGENVPQALKYLKEHYEHVYNKISKCGRQPSYFVVE
FERVDNTIGFVNSDTEHNGSSYMEYRCFDTIRKNASSGPGSGGKSGVLSSGTFIDNEMG
NNNSSAAAASAPAVSAGVSPSLSPFSSDGGDDDDDCSGDDVWGKKMIFNTSGDGSSESSG
QNGGGASTYKFRFCGENTASLSQKENVRLMAMPKGNEDKQLLKNINFLNSALNSVENHV
MCTDENIFDEDDQAEHYTSNKELYKAIVCSNPANVYRVMVELFVNILPRLRNPIVSDIET
VQNLPSNNGSVRTKKMVEHGCTDMRYDIPPYAKGKIRLSAKRACECRKLCCKDVRCFDKSR
EANLTPSQKAGREVEEFPFRNHNHRSNAHDFTFYDKYRARMNKLKDKSKKKVKKIDTFT
TTDDFLLQDRNAFDLLRKCFLSASLHHIFCPDVLVHVRGDSFNINFANNKLECYNERNGI
EEVTSSQTVNAKEALEDITKIKMKRGDDIIDVVKSKGLSLREFSKVKVSKIVRRFNEITNQ
LCNNCNVNSSNGDVDFHVFTSVCVYIHNII PVLEDISIFAELGEELTKLVKECRDVAGED
KTYDDIIRNVEITVKYFKLFNALVKFCHRNYNVAVTSAINRRGYMCMVSNLVGYCKLSD
NAIQYHESLCSLHSSISYADYYTSRNNSEDDGGNSSSEKSNADVAKTMASFYDQFDKSE
DSKKNKNKTSNEILIKMFQMDRVLGMDDDDDDESDSSSSSENEEEEEEEIVKKPAKRRK
VEDVDNKKTLPKPAVKVKQEEDEVMEEVKEAAAEKKEEQEAKEDATEYDDDTTEE
DEKAVASDEDEDEDSKAIF
(SEQ ID NO: 43)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|--|-----------------|------------|
| BAA83091 (BAA83091) HHNBV-XIA | 427 | e-118 |
| Q9YTL7 (Q9YTL7) ORF 48 | 88 | 3e-16 |
| Q18401 (Q18401) COSMID C33G8 | 77 | 7e-13 |
| O35788 (O35788) CYCLIC NUCLEOTIDE-GATED CHANNEL BETA SUBUNIT | 74 | 5e-12 |
| O96127 (O96127) PREDICTED SECRETED PROTEIN | 73 | 8e-12 |
| O96229 (O96229) HYPOTHETICAL 78.6 KD PROTEIN | 73 | 1e-11 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has sequence identity to nucleotides 109 to 282 of CT1048: this corresponds to nucleotides 164907 to 165080 of the genomic reference sequence.

WO 01/38351

67/201

PCT/US00/28888

CT1050

Nucleotide

Genomic coordinates:

Start: 170831

Stop: 171461 (SEQ ID NO: 46)

Amino Acid

MASSSSSPVALSSVASSVMERDEENTLSLRNRNVNKPTPVSAAWVPVDEEDEDREEMRR

LEDFSSDEEDDDNKSCHCDHSDDDDDDEEDPSCFKGFSAGLCSFVRGFFGFLRKSITKKQ

VFLLTSAAVAAIFKTRDVAKTEEGAAATMEENSTDVITGGDGDGIAADVSLASEGEGEN

GSLLESIATTIKTTIENLVDGGEETTEL

(SEQ ID NO: 47)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|--|-----------------|------------|
| O77320 (O77320) PFC0335C PROTEIN | 45 | 4e-04 |
| CAQC_RAT (P51868) CALSEQUESTIN, CARDIAC MUSCLE ISOFORM PRE | 43 | 0.001 |
| O49209 (O49209) PUTATIVE HISTONE DEACETYLASE | 43 | 0.001 |
| Q9YPA9 (Q9YPA9) HYPOTHETICAL 45.2 KD PROTEIN | 43 | 0.002 |
| Q98148 (Q98148) ORF73 HOMOLOG | 43 | 0.002 |
| O40947 (O40947) ORF 73 | 43 | 0.002 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 637 of CT1050: this corresponds to nucleotides 170864 to 171500 of the genomic reference sequence.

CT1052
Nucleotide
Genomic coordinates:
Start: 173177
Stop: 175862 (SEQ ID NO: 48)

Amino Acid
MTRHGVLPVK GRSRHHVILGN VDYTFCTTDN NCVSLDIDFK DNITDQNIQL
LNKKLGKKTAKKIKKEDAPE TKENSDEDIY ATKEFEQTIK GLQTKKGATE
ENAIAAAAAA ATAAAVEKAM LSESEGKSMV INRRARMVLSK RDTSQKQFTA
LKNRESFFSV LIFETGSVIV VGLQDPSLTK LCVIKATTDI ADILQKNISV
ANVSIVNTVS TFNRFHLNFI RLKGFERNK ISYSPNETF PGMFFKLRVP
AKPLLPGETI GEYYTKVAMM RDSKDPNFKM SDWLRIKTAL TFKVVGKITVL
GEGESGCGDV SVVSKLLFGL FHYFMDNNIK MSPKEAQRVR EKYGIPHLEW
YLYIDMLLHS YPYVKPSAEQ VKRAMVDQOH ISEVDRTYYG TKNSMDAAMS
ANLVPKSKEES ISFIKKIRSQ QLFGLHCKPS KETTRRAIDT LSFDPINQDR
WWNKNDQYYG KERCDPFSVA RLVSVSENTN SMMNSRISCQ GKWWLDENEY
KDKLDHIVDL CTEEIVEECE SKGFIASPFLL RKHQKEKIPT PYVLLARACN
QKNGNKMSIN NNSNYLSGSS RAKRNAKLQE KHRVTLARLN TMMASYRFLN
NYISTDIAPD FAKLFGNDVY SLLHMLTNLP KSRGHALTYN ERALSSNEST
YKTPGNAYFS TLFEKSIINN QETANKGNR KRKFSRIGQE KSSFLCNACG
VNLNKGSDIE IKGICTSCDQ NSTSYIENAL SDINRDKKIK RFKAAATHPP
VKQELVDSLSSSSSPSSSSS QTSNKNNRCT PSDFIDYVYK FTDETTGAAPK
VGLVFKMCDI LASLASRRGM EDRPTANYRT SLHSATQNKI NLNKLVSIAI
KETGATETEA QIFNKIIGSE KGLSILCQLV ERRNKDNNVF D
(SEQ ID NO: 49)

Top Blast Hits

Sequences producing significant alignments:

| Score (bits) | E Value |
|-----------------|------------|
|-----------------|------------|

| | | |
|--|----|-------|
| gi 6647869 sp O29874 TF2D_ARCFU TATA-BOX BINDING PROTEIN (TATA-... | 44 | 0.001 |
| gi 6647868 sp O27664 TF2D_METTH TATA-BOX BINDING PROTEIN (TATA-... | 43 | 0.002 |
| gi 3122925 sp Q12731 TF2D_EMENI TRANSCRIPTION INITIATION FACTOR... | 43 | 0.003 |
| gi 135643 sp P13393 TF2D_YEAST TRANSCRIPTION INITIATION FACTOR ... | 41 | 0.010 |
| gi 135627 sp P28148 TF22_ARATH TRANSCRIPTION INITIATION FACTOR ... | 39 | 0.038 |
| gi 135626 sp P28147 TF21_ARATH TRANSCRIPTION INITIATION FACTOR ... | 39 | 0.050 |
| gi 3334374 sp O43133 TF2D_CANAL TRANSCRIPTION INITIATION FACTOR... | 38 | 0.065 |
| gi 135639 sp P17871 TF2D_SCHPO TRANSCRIPTION INITIATION FACTOR ... | 38 | 0.11 |
| gi 2833518 sp Q57930 TF2D_METJA TATA-BOX BINDING PROTEIN (TATA-... | 38 | 0.11 |
| gi 417896 sp P32085 TF2D_CAEEL TRANSCRIPTION INITIATION FACTOR ... | 37 | 0.15 |
| gi 417882 sp Q02879 TF22_WHEAT TRANSCRIPTION INITIATION FACTOR ... | 37 | 0.19 |
| gi 121568 sp P22010 GR78_KLULA 78 KD GLUCOSE-REGULATED PROTEIN ... | 36 | 0.25 |
| gi 3122941 sp P93348 TF2D_TOBAC TRANSCRIPTION INITIATION FACTOR... | 36 | 0.25 |
| gi 3915894 sp P52653 TF2D_ENTHI TRANSCRIPTION INITIATION FACTOR... | 36 | 0.33 |
| gi 1729908 sp P50159 TF22_MAIZE TRANSCRIPTION INITIATION FACTOR... | 36 | 0.33 |
| gi 1351224 sp P48511 TF2D_MESCR TRANSCRIPTION INITIATION FACTOR... | 36 | 0.33 |
| gi 1729907 sp P50158 TF21_MAIZE TRANSCRIPTION INITIATION FACTOR... | 36 | 0.33 |
| gi 135640 sp P26357 TF2D_SOLTU TRANSCRIPTION INITIATION FACTOR ... | 36 | 0.43 |
| gi 1709903 sp P54637 PTP3_DICDI PROTEIN-TYROSINE PHOSPHATASE 3 ... | 36 | 0.43 |
| gi 2833459 sp Q55031 TF2D_SULSH TATA-BOX BINDING PROTEIN (TATA-... | 36 | 0.43 |
| gi 586175 sp P32623 UTR2_YEAST UTR2 PROTEIN (UNKNOWN TRANSCRIPT... | 35 | 0.56 |
| gi 2833446 sp Q52366 TF2D_PYRKO TATA-BOX BINDING PROTEIN (TATA-... | 35 | 0.56 |
| gi 3915739 sp P18428 LBP_HUMAN LIPOPOLYSACCHARIDE-BINDING PROTE... | 35 | 0.56 |
| gi 135634 sp P26354 TF2D_ACACA TRANSCRIPTION INITIATION FACTOR ... | 35 | 0.74 |
| gi 135636 sp P20227 TF2D_DROME TRANSCRIPTION INITIATION FACTOR ... | 35 | 0.74 |
| gi 417102 sp P32103 H1_EUPEU HISTONE H1, MACRONUCLEAR | 35 | 0.74 |
| gi 2833477 sp Q57050 TF2D_PYRFU TATA-BOX BINDING PROTEIN (TATA-... | 35 | 0.74 |
| gi 3041729 sp Q03410 SCP1_RAT SYNAPTONEMAL COMPLEX PROTEIN 1 (S... | 35 | 0.74 |

| | | |
|--|----|------|
| gi 3122946 sp Q42808 TF2D_SOYBN TRANSCRIPTION INITIATION FACTOR... | 35 | 0.74 |
| gi 126745 sp P14873 MAPB_MOUSE MICROTUBULE-ASSOCIATED PROTEIN 1... | 35 | 0.74 |
| gi 4033393 sp P78695 GR78_NEUCR 78 KDA GLUCOSE-REGULATED PROTEI... | 35 | 0.74 |
| gi 126511 sp P12744 LUXB_PHOPO ALKANAL MONOOXYGENASE BETA CHAIN... | 34 | 0.97 |

Comments:

TaqMan Primer/Probe Sets:

5'start=1268

5'stop=1291

3'start=1327

3'stop=1350

5'primer=TTGGACATTTGTGTAAACCTTCAA (residues 1268 to 1291 of SEQ ID NO: 48)

Tm5=57.29

3'primer=CCTGTCTTGGTTTATAGGATCGAA (residues 1327 to 1350 of SEQ ID NO: 48)

38)Tm3=57.45

primerScore=0.74

allele1=

probel=AAC TACTCGACGTGCTAT (residues 1296 to 1313 of SEQ ID NO: 48)

probelstart=1296

probelstop=1313

direction1=Reverse

Tm1=69.02

score1=1.97

length=83

CT1053

Nucleotide

Genomic coordinates:

Start: 175839

Stop: 177107 (SEQ ID NO: 50)

Amino Acid

MSSTDLSKNAFHDWVVSKTDCVFDVHCETDRDCGAACENTYSVDGKEVTKFSCNQSGR
 CARSVYSASSLERANDLGHIIGIIKKNPKEELPESFLWFINHNGGDLFVNKRAAYD
 TMHLSIGKLDNVDTLAQGLDKRMASSLREHLLRKLD SILLQIDKVKEYAKKWILDITQE
 AGTEEDNKEEEDAKKEDQSLSVSEIVDVLGTGTHDPMPLRARGFIQKKIYPLSRNELRELA
 LKELFPEETTSPQVLSRQHDVSTREDLCNESMNAGRAESIFSDPDSGEYVATCACLYSEY
 LTGPACKHKITYRYVIDYDKWKRTGRPEFLTDPVLHFKA EAVCKSTNPNLRAIYSPDNKG
 FLCAPVAELVKTALTFRGSHEPSLIVERDINQAENLPSNSFGVNWPYVNLNRIQDQYT
 (SEQ ID NO: 51)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q26648 (Q26648) TEKTIN B1 | 37 | 0.30 |
| RRPP_VSVSJ (P03520) RNA POLYMERASE ALPHA SUBUNIT (EC 2.7.7. | 36 | 0.52 |
| RBB1_HUMAN (P29374) RETINOBLASTOMA BINDING PROTEIN 1 (RBBP- | 36 | 0.68 |
| CENE_HUMAN (Q02224) CENTROMERIC PROTEIN E (CENP-E PROTEIN) | 35 | 0.89 |
| RRPP_VSVIM (P04880) RNA POLYMERASE ALPHA SUBUNIT (EC 2.7.7. | 35 | 1.2 |
| Q89487 (Q89487) PHOSPHOPROTEIN | 35 | 1.2 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 1109 of CT1053: this
 corresponds to nucleotides 175933 to 177041 of the genomic reference sequence.

CT1054

Nucleotide

Genomic coordinates:

Start: 177123

Stop: 178524 (SEQ ID NO: 52)

Amino Acid

MSASLILDEYLKKTASAVLDVADSFEKIKGEIQSPPEAAALSVALYGAPPKPSASAVASII
 TGERTSLNDKYLSDNVLLKMSVARVGQENNRKRADQAADEIRTIMEDITGSLSGAYRQYSP
 LEEENKVHIGIMNNKTPSIVCGYYTMDTSISSEPLSLTDFQNPTVIANVTKRMESIFSKVD
 SARSTRFDAFVNGVANMMDIKSSIDWANMVENVIKLPDSTPNPCSVDTIVSRDASVVKTAV
 NDIYASVGKSYCREPATQLTFMSEIEKLRKAAVVCFEALMSDTRERAFVEFLFYVSFKEDAS
 NTNSKLFVQNKLSMSGNPRQPIKLVRRSAEETLFGLCFMFKVMPPEFMNCIFNFPTIPHS
 TQYHGLYGTCLTPLLRKYGSSFEKSWAHFEEILSERANAVKKFGVNDTRIDCLDAVANLTG
 PVYVLILDVRLTSLAQRSCTKFLREIKENYLLWNRFSX
 (SEQ ID NO: 53)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|--|-----------------|------------|
| gi 126194 sp P17279 LEU2_RHIRA 3-ISOPROPYLMALATE DEHYDRATASE (I... | 33 | 1.4 |
| gi 6685598 sp O95613 KEND_HUMAN KENDRIN (KIAA0402) | 33 | 1.4 |
| gi 2497227 sp Q04893 YM96_YEAST HYPOTHETICAL 113.1 KD PROTEIN I... | 33 | 1.4 |
| gi 134393 sp P17065 SEC2_YEAST PROTEIN TRANSPORT PROTEIN SEC2 | 32 | 2.4 |
| gi 547908 sp Q02455 MLP1_YEAST MYOSIN-LIKE PROTEIN MLP1 | 32 | 2.4 |
| gi 2496893 sp Q09462 YQ52_CAEEL HYPOTHETICAL 30.9 KD PROTEIN C1... | 31 | 4.2 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 899 of CT1054: this corresponds to nucleotides 177561 to 178459 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=678

5'stop=701

3'start=754

3'stop=777

5'primer=CACCTAACCCCTTGTTCAAGTTGACA (residues 678 to 701 of SEQ ID NO: 52)

Tm5=59.08

3'primer=CAATAAGATTTTCCAACAGAAGCG (residues 754 to 777 of SEQ ID NO: 52)

Tm3=58.08

probel=TATTGTGTCCAGAGACGC (residues 703 to 720 of SEQ ID NO: 52)

probelstart=703

probelStop=720

direction1=Forward

Tm1=68.92

score1=1.92

length=100

CT1055
 Nucleotide
 Genomic coordinates:
 Start: 178529
 Stop: 179348 (SEQ ID NO: 54)

Amino Acid
 MAQTSKMGNTNKRCEEEEEVEEERQQPFTKSKSEPPSFEDKSSSTSSKKKSKSNKHTKTKE
 EQLEFVKDLERSDPTVPDEKVKQVEEEKSPEAIAEIFSMFGIAQDSKFKSLPIERIKS
 ITTKIVIDAINQPVKMLVDHLYHFKEMQNVVEKYKDDSEKLSVILKSKKSPKEFDLSF
 SDYVDRNLNRILVGVIKRVAGAIESKELLQSNMIMNSVLGTVVSNIPYNMKINICVFLTN
 FICTFANDDLTYTFFRDDEKFMVSQVTRYISKD
 (SEQ ID NO: 55)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| 001761 (O01761) C. ELEGANS UNC-89 (GB:U33058) (NID:G1160355 | 45 | 8e-04 |
| Q17362 (Q17362) UNC-89 | 45 | 8e-04 |
| Q17595 (Q17595) SIMILARITY TO MYOSIN HEAVY CHAIN | 43 | 0.003 |
| O31329 (O31329) ERPM | 42 | 0.004 |
| FKB3_YEAST (P38911) FK506-BINDING NUCLEAR PROTEIN (PEPTIDYL | 41 | 0.007 |
| AAD55361 (AAD55361) XNP-1 | 41 | 0.009 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 661 of CT1055: this corresponds to nucleotides 178612 to 179272 of the genomic reference sequence. Hit to public sequence gi|6856160|gb|AF173992.1 to CT nucleotides 647 to 819 of nucleotides 910 to 738 of the public sequence with a 100% homology, a score of 343 and an Evalue of 6e-97.

TaqMan Primer/Probe Sets:

5'start=269
 5'stop=289
 3'start=340
 3'stop=360
 5'primer=CCCCTGAAGCTATTGCTGAAA (residues 269 to 289 of SEQ ID NO: 54)
 Tm5=58.07
 3'primer=GCTCTTTATGCGTTCAATGGG (residues 340 to 360 of SEQ ID NO: 54)
 Tm3=58.32
 probel=AAGTTCAAGAGCCTTCTT (residues 322 to 339 of SEQ ID NO: 54)
 probelstart=322
 probelstop=339
 direction1=Forward
 Tm1=69.00
 score1=1.88
 length=92

CT1056

Nucleotide

Genomic coordinates:

Start: 185432

Stop: 186830 (SEQ ID NO: 56)

Amino Acid

MAVG DYLSMSSVGEATLVGF MILNF INFV TILSL IYAVTDVYRRCKRPSTNGYSGCTTN
 VVSSTLQEANLVTEKDKPVQFVRGLVPRKMMKEYRSDLS PKNVGEYILPSEKETDKLKS
 DYKKGKKVGLLTALSN GHDSNKRIIGPRDLISRDDVKDKSYVFKRLSKDPLVYYSSATSK
 YVRKFSPFRAKKFMTSTQLGSKLVYPHP IRYGTAFVLPTGYVINKAYGMDNEDLHTWNP
 SSSVLVPDSNNDRLTVECAKTDPT HRIGIYGFGGSDDNRRAKEEGYVEMLLCNC DNHKDL
 LKAPLITEYSTNPTEIQVDVAAKRVLF PAPGSEPVKSSQVTSAAHQLDGATGEH DISHEP
 VKLSDTGDYAVGSPIVFKPVYGTSLVNL PETGSPLALNCPCTDKADGIYQVNQKGGILYR
 DMVG YLNANPVEAASLSSSDSSSWLTTGNKISSVTCEGEKIKKIV
 (SEQ ID NO: 57)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q99175 (Q99175) HYPOTHETICAL 58.5 KD PROTEIN | 34 | 2.2 |
| CAB49723 (CAB49723) HYPOTHETICAL 52.6 KD PROTEIN | 34 | 2.9 |
| Q9Y6X0 (Q9Y6X0) SET-BINDING PROTEIN (SEB) | 34 | 2.9 |
| Q9ZXH3 (Q9ZXH3) INT44 | 33 | 3.8 |
| Q55105 (Q55105) MULTIPLE LIGAND-BINDING PROTEIN 1 PRECURSOR | 33 | 5.0 |
| Q9WXH8 (Q9WXH8) PYRUVATE ORTHOPHOSPHATE DIKINASE | 33 | 5.0 |

Comments:

TaqMan Primer/Probe Sets:

5'start=663

5'stop=684

3'start=705

3'stop=725

5'primer=CGTGATCAACAAAGCATACGGA (residues 663 to 684 of SEQ ID NO: 56)

Tm5=59.59

3'primer=GAAGAGGGTGGGTTCCAAGTG (residues 705 to 725 of SEQ ID NO: 56)

Tm3=59.41

probel=GGATAATGAGGATCTACA (residues 687 to 704 of SEQ ID NO: 56)

probelstart=687

probelstop=704

direction1=Reverse

Tm1=68.99

score1=1.88

length=63

CT1057
Nucleotide
Genomic coordinates:
Start: 190875
Stop: 193236 (SEQ ID NO: 58)

Amino Acid
MEYMEEGDIAERRSEGVYILDENSACVVNVKSIRNRLGAMDAEEAQYAQDISAQLVTHI
IRLAHCSESNIKIDTIASIAGLFINNIFDNNSTKNKLKTYNQFKAESQNKSSVLNIFGSL
DPLSMLSSFMGSDPAKSGGENLDKSLGVLFEVLQYNPCKIDDIVLLEMCPKCAACTGL
KEAIRQEQPMEAMLLFFKCINHNRFNFGSDIKSAYASETCMRYSQDERAVVPLRSILLG
CLDRDDPAHTLSSFGDTIEYADSDNAWVSSLFAAVSRMPMVDRAVIAHFYVYTMLSRHRR
VSGDSFKQFVYTVFVRMIYSAIEILFCDTENSSVECDGKHFLSYVNMVNVSVLGSTFNV
LKAYRSWVVDQASVAPVLDIISGGWKKNYSPDHIKRVAYDISQVINHLASPSRMVKGNN
KASNVTSGLDIRSVRQAEKYIPFGILENKAGYGVINIAKHNISRPAREQSNGRNFNCNA
LHILPSIKGCEALGAQKGSADQTVNVFDNFVASHMDIAMKKQSGKILGLLTSMIDRQGL
TTSFPSSEAEYKKRIHDFTRYVIFSSTPINDELVNSRCILPHSNVLNSPISLRNIDPESV
PDRFRHFLLMWQRPNIDEPNLSALTTSQLELLLSKNQKWDKLTTRAFFNIDRINFQAD
AIIKNVSGSGFLDGSKTASSSSAPNFFQIFSGAECTAKQLQSIRKFIGESMOHVQKEWS
SAVNNNGRNGVENYDGLNAQFSEELFELLYKLIIEEDMRPSSLIASSEFLSNYVNAMELL
IRANAS
(SEQ ID NO: 59)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| CAB52133 (CAB52133) RCOP C7 (FRAGMENT) | 34 | 3.0 |
| IDHC_SOLTU (P50217) ISOCITRATE DEHYDROGENASE [NADP] (EC 1.1 | 34 | 5.1 |
| Q23615 (Q23615) ZK822.4 PROTEIN | 34 | 5.1 |
| YSW1_CAEL (Q10017) HYPOTHETICAL 63.8 KD PROTEIN T25D10.1 I | 33 | 8.8 |

Comments:

EST confirmation of the predicted transcript:
An isolated EST has equence identity to nucleotides 1 to 636 of CT1057: this
corresponds to nucleotides 192639 to 193274 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=1214
5'stop=1236
3'start=1281
3'stop=1303
5'primer=TCATCAATCATCTTGCATCACCT (residues 1214 to 1236 of SEQ ID NO: 58)
Tm5=58.13
3'primer=CAGACCTGATACTATCCAGGCCA (residues 1281 to 1303 of SEQ ID NO: 58)
Tm3=59.10
probel=AGGCTAGCAACGTTACAT (residues 1262 to 1279 of SEQ ID NO: 58)
probelstart=1262
probelstop=1279
direction1=Forward
Tm1=69.01
score1=1.98
length=90

CT1058
 Nucleotide
 Genomic coordinates:
 Start: 209615
 Stop: 227849 (SEQ ID NO: 60)

Amino Acid

MDQYPEVRDTPQTEQEQAQAQQAATTTAAAAAAAPTQYSNTVSAETLSAISEDGKLE
 RSIAASCWINNLPDEKMAQRVQFHPLSSTTTYDSENVNPGSSVFLKPRALPTGGTCLA
 PNYIAVPTLRAASEIIDSIASSTSLYQCSMFNSWNLPIFMSNSKHSQFGDRVIKRSMIRN
 CFSKQKNVENLLKELRRRKVNAAKAFSHAVQOKSAVNTALAAWNAGSAANLEKLVDFCKL
 KYSPPDRKYKAGGLFSASATAQSQSGTSSSSVEHTSNDFLLDILKRHKGTSLDLSATNTF
 DTALSRVFTEFKEQARAADDAADSDHLSASDPFISIVRHNSRREGILDSPVNI GMLAPR
 SKYSVAEYLMADRDESADIAAKIGTKIATDFEALRGDNNKRRADTSVDDLKESLADSIEK
 TSIKNTGDI NSVTNIPTDTEEYEFSLHITQLFAQAFLETMGSLLSCAFGVQFPFSDGEGA
 ALERIIRKTDPTDGKVSSEMDPSSLSDQYLLLVGNFQVSPFHVSDPKDIVFGRQVTPNTP
 LSIITRSKNKNETSTIINFRDRLLVNDTVLRDATQNVSTSTPSQRRVPTAAGEPKKPM
 SGCLPIIRGPQVVTRESDDMISGLVGDWYISLGVYAMGSSAAAIAAGHQRALASAESIN
 SPMKKKFSKKGKYTEEEKRIKKAMRRNADRSARILALLGQTDQYGYVEHNSTLDSFWS
 SNAAIRAKAKEDALSRAEILAVRKQLDGKCSSSRDEYSMVERYLRDSFFRSVNRSGGGYE
 MFDQGFDMGRFADFLSDNSAARNWQOYAEVMRGLSKHEKRVFNIEGLFSALNSFKFPLV
 PEQGRKKTVGGRHRLNNLKAANKIINGITEMTLQSAIDGTGISDIIGSVSDGWGNTTAQP
 SRVKALKTSLNFSGNGNVVSI PVSRAVKCAAGSRGGETLKCVDIPSVIIANLISDKRILD
 QLCGGGMNLAHEITNFIETIAGKEHTGKESVFLSPRLSVILLRYIWFNAVVSLTDSNIK
 MPLNTMSEGTGDDIYRDYLAIRGMVNNYNSSLSSISVKAISDRYNCGSGNTSTSNKNVTI
 KTQGELLTVLQQTANALSFTNKGVGATPDANMANVISPIANADVVKNTNVVVSGLDR
 ITETINFFSFLSQIKTMNENIEEYLRRYRLGEGLDKKELDNFVYPNIAAIVKRELGVSGS
 ALSSNLDTRPITIDLNTEQPLIVKASKGYASNRYAKLENKTRTAAEQAQMEQYNAQMA
 ANTIPQLVNRLTIPGSITADTAINVVKAFTENGFEFSNAETHLGVMGNAINEMQPLFTDGF
 NVANKRLTVNVGVSVKLIQNGLTVSLILAHSKASPYVFKPLVQDFAKLLAVTAETSLV
 SRSQKSFPIPPSVFSSGGLFKIDREMFDMKTDYVVEVIRQLSKNATAIERCNDSDSA
 ARIAKSGEIYNKDVASTTAAPGTSSSALTFLANNLQNPAAKVMGALPHFDMAVVPKLHG
 ISHDQMFRLSTYYQGIHKMELNSDCKPEEWDNSLPGNRASKFFGLSSVSDNNRSFNLALD
 TLLASPAEICDLVTREMVKTSNDIVHNI GSNSTDALQKSLQVGASAVEKYDESTLSTKE
 TDVYSLVSLAKSKSPLSSSSLSSEGLTSKEIDRTWNTPALLGTAKTTSYSVSEDALN
 APLSAVLDFRRNVVDATKSLYEVAAVCSVMKEEDVRSSSRKIMGVMEQESVPMQDIGID
 RIASLVSTVATPKQHRRFLQTVNDYKNYLIRKVASNPLSSRLGGISPTSGNTDYNLKA
 YDGVVSSSSMTFSSMSVSDRFWSGVFSQCLETGSPMFADAGHGGSNMFQITAPKLYGSR
 VNTYAALSSGVERLDSISSATQERKNRIAKSIEALETFVTDVVGGDTLDQLRKAQNM
 KLSDIATSNISYSDFGNIDCAKIMKNVTSKMTARQOSTILSLLHELGLVHKQPPQLA
 TQFALASHVIKAKYVTNDLNNIHEKETFSQLMAVAGVADYYNVSAAMCQRLVASDVTMF
 LGGTMLQQGLFVSFLLNNVLSQVSDNIKMNELNDETSKLLVKLVGFCGTVSDALGSRHV
 SSIRRVQNEEDKKLDRSFVTSLSAYRDLRKKTELYRETDTINKLFGHQNFMSYESSMLK
 RTSLVHDAVSGPRPRRYSTLEDVLEAPSTVHKSFMVSYPERAAASRRVKRAGLRALADNR
 MESLYGEEVLNMRSSAVSSEMMDIEYEGGGFMMMSIDEDDDIAFIDSEESSESSTDFSS
 SDEYSDDSEYDFDDDDNNGQSPYSTTSYSDALDRLNSAAKPLTAIYGCRGEGEDDEEND
 LYEEEQERRRRSSKMGKILRDLHESDDDDDDYFDDEFDGERSMSETIATRRAGRIQYGP
 FGLSHSNILNRPAAKARAFLTRGKKFRPSAYDRFFMEDDDSLFSDSTSSSSSDSPFSS
 FSKGRKCKRRTSEDQCAFVKRVRAVFPTRVTMINGRVSMITPVTSENTVGFYENYQAN
 KRERARLIEYKIVKGASATLPDEYVEGRASKQVSPRELRRSLIKAAAYVARTQESNLNI
 IFDALTTTSNATLVNDPSTLLGDTLLFAKQLEAITERRNRLMKDLTEISPSLFTSFGDAS
 KDTQMMAADAKQIVSGGNFKSAGYLGVPRLRTLASCIGKTNTVDRLLATKNKNHLEWMTAA
 IVFARSFNDTTFHALEDTLKMTSALTDMSAFTNLVGSEHSQRLKVKSTLLDSIFNTRMA
 HTEAVMGLVYPTAFINHEMPSDYTORREMQSLALNIRGVNCSQLPRKDIDGTAGLLTFI
 TSRKFAGYGGGERGGLSLYRMSIVDALSCPSDNRLKGAVSLEVGKWQDMGEEIFYKRSNDL
 VDSCSKNNISLENAVGPPIARFVPNGTNMADIGMTDII SRTVKDDASMI RLRRAEAGAGAA
 GKFITASAMGNLYGGIDTVVNLTEKLYDSFVLLQDSDFNTPTMATAIINRMKSRKHKA
 LKTPFGGDIATYKNFPSSSEAI VVRAKEMRNSISTIVMDISKSRGINSFSSRSGSTLAKI
 STSEFERILETSAVLSNTKANLRTIENRLAEHYNKLKQFHSISNDGLSETRAVVAVIAES
 LTPVYADDTSERGASVSELLTDNTLLKFIVQNELKNI EAKRHVTAIEGSSQLHEKMLS
 LLVASADINRMSAQNNLECKKLEGN SNFVPM TNDQGGTFIKHKETGIWLKTDENNTSS
 IKDNDQRRVAKTILAIVEDNRNATIRSRLQSLCFGKYAMNDIFALDDADIKNMDKLIKEL

GEALAEKASPPSSSAISSSSSSNTTSSSSSPSSSPSSSSSSSMDYSNNLAKTIPYMPIVF
 QNKQSNVNSSDASSSSPSSSSSSSANIDNVEHKKVALQQLQTQESNDLSNVLSVTTKHRF
 ASHNQAATVGIFNGRQHAETVVAIPNANKANNNATVSAGQGILTRFSAPENVSSSTSMQLP
 PSSSSSSNGDDNKVPVTVRLNQYANSILSSIENASEFKDLKEAERKIDLAIQAASTETK
 EMVTVSKCPSANQTAITAISQAKSLKKSALLELLEKVIKAVEVYTPDSSIAAVSLPVNGDS
 MVSSSSSGSGSAPSSSSSSSSSSSSSSNVDYFNYAYGKLKNIDENTEEGAETVQKNMVEQD
 AAVRIPLLVSYAPFSEMMRRAIDKLNEYQLIDAICTKIVSDTKQASSWAIKETDKELDM
 DKEQVISKINNLOQNFESNESDKIKMAISVLNKRNELELQNNKTRSFIEETKSRIEAGGG
 DVANFKEIIDYENTSENDNNLFQSLKAFADNSGTVYTPDMSNGRDTKSDSKFVDMYNK
 QILEGGIKLINEGQNTVKVDFSKALEAFPRQSNGASEPVSSSVERRQRERLQAVEMFMA
 IMMERTESLRKRLADSAQWNTVNNVEETVNSGMVNIKSERLTEIRNQAQIAESTALNSI
 NDEIVESPLTSLGARDVQLLIKVDRVGSIQQQQQQQQQQQQLPKLTATEQRKEQQYAAD
 RVVYDPSYTCFLQPLHETIKRISSVYNSKNKGPLSNTRGVPTSDADLQMTITDLSRSVL
 DSSSTSSKKMLYENVPSIIVPGLCQQCAMMITNVHEATHTSPHSFNFENKRSKQLTEML
 NAATSSSDGPAVRHDVLTMLSNNGYVKDFGFTHRQKVACITPVNTLLGGTFSGNVAPNT
 VILPTSELFNCPGVENDKFRSMVNRTTDKNVADAPKSSASIVETLARTSPNAEHLYFPFK
 DQRRHFNSITDAIISGMSGESSQLNTTCDQNLVNI DOTTGFPVFTGRKQGERRIVHTEN
 TMEGARKDKNSGIPSC TKDRQTYIDMGTKFEMVAPGSLLNANKEETLRLNRLSDINNVRHY
 GTDVHVAGANSARWIGEVVRAASSFPDGDKESAMKMLLLGSVAISAQKSASHINDPTA
 LLSTNTSIQNLVKEAFPDPVCSSNYLGSAAESTFATQLAYRQRLFPNGDDENVTVSNICP
 MDLMGSTKRYNDAFNNIFGSKMTSTNKKGSNCENLLKSAMSNVPAINTAFGAFAEEASSSV
 RNRLSPLYEDSTKYSSNQLAVQAMTDTAVDALSAVSTVVGRQNGRNTLLSLPTSITSIAT
 SGRPSLSYSSDMKSNLIKTI SRINRDASLLSMGDSQVAAGSSFFNSFLRSSIPVTTSDQ
 GNVAAAEIVLGTILDKTVEINKRFEMLGGGKMVAGSPEARAIQRNTMSSI LQMNENELAR
 DLCEIENKIETRQLRDAFQDLKRSMLMTPGGVGAISSGASTNNVPLSLLMSRVDASSGLL
 MNNSANVMEAVDSFNTTPLLVRHMMLD SGKSPVPMKEIRSMLTQPRALTARALLSESS
 PLLTEICLYNTRDTQPERAVDRLLTSAYLVKQAKRFDGVDPAFPAAALTCASHMLSSMDS
 HTKSSFMNLIKLMHTDQCFFKNIERFEKFLGRYGYDEYAMSHKQNCNCPFHLHHTFTPSD
 NEHLVSSFAFARPEVSMEEIRATPYQANKLISDKHYVMNMSKIDSRVTGSSLLKKVSEWT
 EMRMNSNFNGTFEPSRLALSNSGMTTAGVNLDVIVKPNNARSVLGILECHRQHVCTADAK
 GTVASAMPVAFQATDGNNGNESELIONALPRNRYIQKSTMNAQTUVVFANVLEQLIADLGKV
 IVNELAGTIAESVPESVYENTKEMIDRLGSDDLFSNNNGGVESMDYEDSETTSNNGPVL
 ISEAMKNAVYHTLISGKAARPENVPFASCASGPLAFDFLLSKGDTFEEKNAEQGAAAASVS
 STYSSSSNTTLRKHLARVFEAISKQVTD AEFKDILNDIERNISSDYTNCPNTNQNAFAL
 AIKREFSRIVSFLTLRKNTIPALVDPKGALHEKVAIYLTLLSTKSKLENFFQYGLSNSS
 SVDLSHLKPINCNNVKNIEDTFMYRNVHPILIMALPENFTALLQEQMDPDTAIESRRS
 LTTFLNHPNTASMANARAAVAGGGGNPMGLYLSSHILHESTVTTSNPVTDTTENVNYS
 SVTQDPVMVVPFKDSARLIVNNNTGIDVLNDKSCNYLQVSMPSSESSGLVTNTGCSSSS
 SSSSSDTFKYVRRDNTPVNLPRVTPAVLCSDASSNLLDVFSRADIVLENMNVRFGEFMEI
 IAAVSFKFGLTKEEVIKOMVSONNNNNNNNGKKTVDPTGDIVITNATFPDTRP
 LYTAANGGTSSFKWGDINDRKMHAKAFPTFFIGNPTAAATANGVPLTSEGISLTEEKRRK
 IAGISEGSIGTGALRAAANTRLSSDMEPVMKGWNNIVQLQQTFFKKASDKLTHLLRSGGIP
 PRSQETNAIINKMHSFKTLEECRRVIQDEAALLVATSDLLTGGYGGDAALAMVSPVRPE
 MTGLIGAISAPVRGISHLLKLGVSANAAIRKRLNLPSTNGKTLPEHGIVHKSATLLL
 OSDSISNLYNTDLQDVVSNARDNNNLGRIMQSLGLKGNAGDLVYSARQLTDLITVPEYG
 NNRDLTKRQAILKMLISNPEILENVADTIYLTGKNALAPVSAQEMACASLTVGGSGGGK
 LSSDDNVQSLNRLYFRV
 (SEQ ID NO: 61)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| SR40_YEAST (P32583) SUPPRESSOR PROTEIN SRP40 | 80 | 5e-13 |
| O94317 (O94317) SERINE-RICH PROTEIN | 80 | 5e-13 |
| O95815 (O95815) DENTIN PHOSPHORYN (FRAGMENT) | 80 | 6e-13 |
| YM96_YEAST (Q04893) HYPOTHETICAL 113.1 KD PROTEIN IN PRE5-F | 76 | 9e-12 |
| AGA1_YEAST (P32323) A-AGGLUTININ ATTACHMENT SUBUNIT PRECURS | 73 | 8e-11 |
| Q9Y076 (Q9Y076) PROTEOPHOSPHOGLYCAN PRECURSOR (FRAGMENT) | 71 | 3e-10 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 704 of CT1058: this corresponds to nucleotides 227240 to 227943 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=8829

5'stop=8848

3'start=8894

3'stop=8917

5'primer=CATTACAGCCTCAGCCATGG (residues 8829 to 8848 of SEQ ID NO: 60)

Tm5=57.78

3'primer=GCAGAACGAACGAGTCGTATAGTT (residues 8894 to 8917 of SEQ ID NO: 60)

Tm3=57.76

probel=GAGGTATTGATACCGTTG (residues 8861 to 8878 of SEQ ID NO: 60)

probelstart=8861

probelStop=8878

direction1=Reverse

Tm1=69.00

score1=1.99

length=89

CT1059
 Nucleotide
 Genomic coordinates:
 Start: 228374
 Stop: 230564 (SEQ ID NO: 62)

Amino Acid
 MDKVCVISNTRERTFKVPADLLCVATEPEISTKEEDAGIEIETRVVVF SRCVSVQELHTI
 NPNDGFSVQLFKDY LKLQSAQGKKPIGLYIQIKAGEDLERRLISGGTAYLDPATHLFYL
 DFSLYPNYSIFNDISSRLKIIDEDTYNGVVF SNSEEKEDALVLRVTFSTHEKAIEAAI
 KKIMLRKVFFKDGDLDFGYLRIPKSKLDKFTPYFRSQYGIVNVEKNIPGYIWGEIMKQRV
 RCSRWYLYNTDSEWEYKNVAEERVGPRQLVKKYGAKCENLCFRDIDLKKEAKEKRDIER
 ETESRYVVVTLTHKHEMPENMPYFGPKCSVVRLETRILLCFVDEISYNDEDVDEILSEN
 RSLRNV SIRHKENVPVHTLLKKGVS IHARFTLNLGLDDALIILKRIPKTYFEDEELQAACA
 HVNLEQYEWLCSNNRGNKVEHVKS RVVTRAVKRRRKCRHWIYFDKDTLNLNYKYFDKKVT
 ASMASKICNAKHDCLVFHRKMELEDLTESAYFKVEPSPINFAKLKSCPDVKYVQKKT DGT
 FSVIRFFRNMTKGD LIQRMDLFCRFIPDSHTITLLSRADFYACKRGESMHMCTNKHRI LH
 YKFSNAPHAAIEQITNIISDTRGRKGIHIEYAIENVQEMYEEDGRRYEAKYTGTLTEYKR
 NEDKTFKSLLAPHLTPVNKPYNINHLYEQYGNFDEELEDKLRSGFISYDTYVTAKDNWGR
 CATGKGACI
 (SEQ ID NO: 63)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| YO61_CAEEL (P34600) HYPOTHETICAL 84.7 KD PROTEIN ZK1098.1 I | 36 | 0.95 |
| Q9ZAJ8 (Q9ZAJ8) BONT PROTEIN | 36 | 1.2 |
| Q01794 (Q01794) MAJOR SURFACE ANTIGEN MSG1 (FRAGMENT) | 36 | 1.2 |
| Q9X708 (Q9X708) BOTULINUM NEUROTOXIN TYPE B (FRAGMENT) | 35 | 1.6 |
| GLND_HAEIN (P43919) [PROTEIN-PII] URIDYLTRANSFERASE (EC 2 | 35 | 2.1 |
| BXB_CLOBO (P10844) BOTULINUM NEUROTOXIN TYPE B PRECURSOR (E | 35 | 2.1 |

CT1060
 Nucleotide
 Genomic coordinates:
 Start: 230616
 Stop: 231582 (SEQ ID NO: 64)

Amino Acid
 MCTLKTYKMTTSTEISKNLSDVLSIKATGDWCSNIKTVFSPFTEGKGNLPSLPFTRSPN
 TTCGSREAANATEHFITVFAKDKYERKRVKRTIGFTLDNTKELTPNRYLVADVYSWQEEK
 MVFEGFCVPPGKSGTFVRYSNEDKSFLADTGRYMKKKYDDPENKTSSGGDDDDDDDDDD
 DDNNNVVDVYEENDPRNVFEVEKDEKYACTFSILVYRAMKKSPVCRGLLVETDGPSSHPK
 RAPSAFNPFGGSSMLNGYGAGADALEEEDVDGVPERERITNFALKRGPATGQNFVSVKL
 EHDGSKADLYNVTCSKQRGV
 (SEQ ID NO: 65)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| 097300 (097300) PFC1035W PROTEIN | 51 | 1e-05 |
| KEX1_YEAST (P09620) CARBOXYPEPTIDASE KEX1 PRECURSOR (EC 3.4 | 47 | 2e-04 |
| Q83970 (Q83970) (CPV) | 46 | 3e-04 |
| SIS2_YEAST (P36024) SIS2 PROTEIN (HALOTOLERANCE PROTEIN HAL | 46 | 3e-04 |
| YB00_YEAST (P38114) PUTATIVE 126.9 KD TRANSCRIPTIONAL REGUL | 45 | 6e-04 |
| 077384 (077384) PFC0760C PROTEIN | 45 | 8e-04 |

Comments:

TaqMan Primer/Probe Sets:

5'start=477
 5'stop=500
 3'start=555
 3'stop=578
 5'primer=CGATGATCCAGAAAATAAGACCAG (residues 477 to 500 of SEQ ID NO: 64)
 Tm5=57.85
 3'primer=TCGTTTCTTCATACACGTCAACA (residues 555 to 578 of SEQ ID NO: 64)
 Tm3=58.34
 probe1=ATGACGATGACGACGATG (residues 515 to 532 of SEQ ID NO: 64)
 probe1start=515
 probe1stop=532
 direction1=Forward
 Tm1=69.04
 score1=1.95
 length=102

WO 01/38351

80/201

PCT/US00/28888

CT1061
 Nucleotide
 Genomic coordinates:
 Start: 231602
 Stop: 232799 (SEQ ID NO: 66)

Amino Acid
 MQLILSHHLTMAGRVELVTGPMFAGKSTYLKNIYQQENGGNKHCLFVKHSLETRYGCGTG
 TIVTHAGEVIEGCTTVSSIKELISVLPEVVVDVILIDEGQFFTDLVLVNRLADKGRIVIA
 ALDGTSDQQMFSPHKLPLPYTNSIVKLASKCMICKIDTKEAPFTVRFGNDNDNNVICVGG
 AEMYAAACRDCYKKINKKKKNKGKLVVLEGGDRCGKSTQAKLLLTNKNSPLYGGEYMCFPD
 RSSHTGKLINDYLTKKIELDDHAAHLLFSANRWEVCSKIKQLDDGIHVMDRYYYSGIV
 FSLARGVDTVWCASDEGLPQPDVLVLLMLLDVEKCSNRDTFGVERFETNSIQERARALF
 LDLANKDEKNVWIKVDARGTIEEVQTKIINIVYNIVEE
 (SEQ ID NO: 68)

Top Blast Hits

Sequences producing significant alignments:

| Score (bits) | E Value |
|-----------------|------------|
|-----------------|------------|

| | | |
|---|-----|-------|
| O74528 (O74528) THYMIDYLATE KINASE | 165 | 6e-40 |
| KTHY_HUMAN (P23919) THYMIDYLATE KINASE (EC 2.7.4.9) (DTMP K | 159 | 3e-38 |
| KTHY_MOUSE (P97930) THYMIDYLATE KINASE (EC 2.7.4.9) (DTMP K | 157 | 1e-37 |
| KTHY_CAEEL (Q22018) PROBABLE THYMIDYLATE KINASE (EC 2.7.4.9 | 148 | 8e-35 |
| KTHY_YEAST (P00572) THYMIDYLATE KINASE (EC 2.7.4.9) (DTMP K | 143 | 2e-33 |
| KTHY_SCHPO (P36590) THYMIDYLATE KINASE (EC 2.7.4.9) (DTMP K | 141 | 6e-33 |

CT1062

Nucleotide

Genomic coordinates:

Start: 232848

Stop: 233334 (SEQ ID NO: 68)

Amino Acid

MLPRKTLPDTENG YFVLDES LLEKVYYDNNNELIVRVGGIYMQICKSKYIFHHDDPERFF
 YSVLEDYHPIKEIVERLAEEDGVFLGPWEFLSRKQVNLQHGCYKALLSLPEDKYCNLLLP
 QQMKTNLEKMEEIQRTLHRSRTYNT PQIELSDQLDGCVIC
 (SEQ ID NO: 69)

Top Blast Hits

| Sequences producing significant alignments: | Score | | E |
|---|--------|-----|---|
| | (bits) | | |
| O60678 (O60678) PROTEIN ARGININE N-METHYLTRANSFERASE 3 (FRA | 31 | 5.4 | |
| P95966 (P95966) ORF C04027 | 30 | 9.3 | |
| Q9ZW94 (Q9ZW94) F5A8.4 PROTEIN | 30 | 9.3 | |
| O58387 (O58387) 358AA LONG HYPOTHETICAL PROTEIN | 30 | 9.3 | |

TaqMan Primer/Probe Sets:

5'start=223

5'stop=245

3'start=309

3'stop=330

5'primer=GAACGACTAGCAGAAGAGGATGG (residues 223 to 245 of SEQ ID NO: 68)

Tm5=58.17

3'primer=TGGCAATGACAAAAGAGCTTTG (residues 309 to 330 of SEQ ID NO: 68)

Tm3=58.99

probel=AAGTGAACCTCCAACACG (residues 284 to 301 of SEQ ID NO: 68)

probelstart=284

probelstop=301

direction1=Forward

Tm1=68.99

score1=1.96

length=108

WO 01/38351

82/201

PCT/US00/28888

CT1063
Nucleotide
Genomic coordinates:
Start: 236678
Stop: 238604 (SEQ ID NO:70)

Amino Acid

MVASTPCPGPGVPPTQELLSTNFLEAHKLVVVELLLPSYSSDVVYCDSETYTKPIPIFGNK
SIVSTIGDYVLSNPNEVDVSYQMVSSVLEKFPLLFHCTYKTNEEDKGIPLWKKLYNKRKFK
LLNSLLVHNNKNWTPVPAIPFDRENICDASGRSVMSEIMSTSTFTICKNNTHYLFDM
NMERGKQGGSFLLHFFASRKNSFTNFENEEMDSHVLSNIAKFICNEKEKLDSPANGKIP
CPDKTNDEGYIPLEIAIMEDNYPALLYLVCRYGASWANTYGDHNEKAFKFAIRNDAKDCL
EIEFISDHYSFNKNVTKEEFVKEKTVECVGCLYDIEDEKRCYKLP CGHFMHTFCLSNKC
SKANFRVCVKCFQTFDDTIFRKCPPTIQWKMGINQTTNHKEMDLFNRAFDTYLDFICSYNV
KLDKKS KPKHKPENKKVEEELAKRTAEIEEAIAKKKEEELAKRTAEIEEAIAKKKEEELAKR
TAEIEEAIAKKKEEELSKYNKIIIEKGKRLNEECVKLRDISTAAINMYKEKVRINGVLLK
DSDQELAEAKERLRKILLLEEETKLDRLFPRPKRVEERIFLT KDDETLAFKLAEKKTED
IIAKKNNQKGSERRDGEYTTITSHIEKLPOSTALASVCVLNE
(SEQ ID NO: 72)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q26938 (Q26938) KINETOPLAST-ASSOCIATED PROTEIN (KAP) | 74 | 3e-12 |
| Q9ZIU2 (Q9ZIU2) VIRULENT STRAIN ASSOCIATED LIPOPROTEIN | 73 | 5e-12 |
| O50870 (O50870) HYPOTHETICAL 54.3 KD PROTEIN | 73 | 5e-12 |
| Q9ZU69 (Q9ZU69) PUTATIVE VICILIN STORAGE PROTEIN (GLOBULIN- | 68 | 2e-10 |
| O23230 (O23230) TRICHOHYALIN LIKE PROTEIN | 62 | 1e-08 |
| MNN4_YEAST (P36044) MNN4 PROTEIN | 62 | 1e-08 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 362 to 1 of CT1063: this corresponds to nucleotides 238163 to 238524 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=775
5'stop=797
3'start=850
3'stop=872
5'primer=GAAGACAATTACCCTGCATTGCT (residues 775 to 797 of SEQ ID NO: 70)
Tm5=58.33
3'primer=GCAAACGCTTTGAGAGATTCATT (residues 850 to 872 of SEQ ID NO: 70)
Tm3=58.56
probe1=TAGGTATGGAGCATCTTG (residues 810 to 827 of SEQ ID NO: 70)
probe1start=810
probe1Stop=827
direction1=Forward
Tm1=68.95
score1=1.95
length=98

CT1064
Nucleotide
Genomic coordinates:
Start: 238658
Stop: 239438 (SEQ ID NO: 72)

Amino Acid
MSTCSNLLSVFGGDWTTTFFDLVHTRQECDKKREQDYSFFITETCKGENIGIHSYEHT
SKIIDTGNDSTSIEEVLNIYKAINHLENILKLNKGEKIILMDVETMILETHKILMKG
ILPKGKNGSFSTCVRFAVNKNNERHYYPVFETEKEAFNSIQNLVDYYNEIVAHTNDQIKI
IKACAYFMYNFLTILHPFNDGNGRTARLLYSFLLKGNIGIVPHFSPITHPRDQFVDTLVYFR
EHGDGRPLLYVLLESIKNK
(SEQ ID NO: 73)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q23544 (Q23544) ZK593.8 PROTEIN | 50 | 2e-05 |
| Q9ZHQ9 (Q9ZHQ9) HYPOTHETICAL 26.1 KD PROTEIN (FRAGMENT) | 44 | 0.001 |
| O68899 (O68899) HYPOTHETICAL 29.5 KD PROTEIN | 43 | 0.002 |
| Q48249 (Q48249) PLASMID PHPM180, COMPLETE SEQUENCE | 43 | 0.002 |
| O75406 (O75406) HUNTINGTIN INTERACTING PROTEIN HYPE (FRAGME | 42 | 0.004 |
| AAC96089 (AAC96089) HYPOTHETICAL 23.9 KD PROTEIN (FRAGMENT) | 42 | 0.004 |

Comments:

EST confirmation of the predicted transcript:

Nucleotides 1 to 677 of CT1064: this corresponds to nucleotides 238717 to 239393 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=397
5'stop=419
3'start=450
3'stop=472
5'primer=TGCGTACGCTTTGCTGTAAATAA (residues 397 to 419 of SEQ ID NO: 72)
Tm5=58.67
3'primer=TGAACGCTTCTTTCTCTGTTTCA (residues 450 to 472 of SEQ ID NO: 72)
Tm3=57.92
probel=AATGAACGGCATTACTAC (residues 424 to 441 of SEQ ID NO: 72)
probelstart=424
probelstop=441
direction1=Forward
Tm1=68.96
score1=1.88
length=76

WO 01/38351

84/201

PCT/US00/28888

CT1065
 Nucleotide
 Genomic coordinates:
 Start: 240712
 Stop: 241192 (SEQ ID NO: 74)

Amino Acid
 MEDLKSTIERVYEERVENLEQWTNTVEEEERTVSAIDSVLEEQKRALDAWEAAIKEREND
 LAVKEGISALVFNAADAKTRKELINTWIAERETSEKRRKEATSTNNQLKNQMSSLVNTTK
 TLKEKYNKYRRSAILNMQYINNKRDEASQFWVYTNA
 (SEQ ID NO: 75)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| RADI_HUMAN (P35241) RADIXIN | 43 | 0.001 |
| RADI_PIG (P26044) RADIXIN (MOESIN B) | 43 | 0.001 |
| O40947 (O40947) ORF 73 | 42 | 0.002 |
| SSP5_STRGN (P16952) AGGLUTININ RECEPTOR PRECURSOR | 41 | 0.005 |
| Q9Y489 (Q9Y489) CENTRIOLE ASSOCIATED PROTEIN CEP110 | 41 | 0.005 |
| Q21952 (Q21952) SIMILAR TO MYOSIN HEAVY CHAIN | 41 | 0.005 |

Comments:

TaqMan Primer/Probe Sets:

5'start=228
 5'stop=249
 3'start=297
 3'stop=320
 5'primer=CGCCAAAACACGTAAAGAATTG (residues 228 to 249 of SEQ ID NO:74)
 Tm5=58.34
 3'primer=TGATTATTGGTAGAGGTTGCTTCC (residues 297 to 320 of SEQ ID NO:74)
 Tm3=57.45
 probe1=AATACGTGGATAGCCGAA (residues 253 to 270 of SEQ ID NO:74)
 probe1start=253
 probe1stop=270
 direction1=Reverse
 Tm1=68.84
 score1=1.84
 length=93

WO 01/38351

85/201

PCT/US00/28888

CT1066

Nucleotide

Genomic coordinates:

Start: 241184

Stop: 241505 (SEQ ID NO: 76)

Amino Acid

MHKFSNKFYFIKGVLIIFVDPDVFSIFLLPPLGVRHKNGGGGNEEQKSGPSQKHHIPG

PVLIFVLIIVIVGSVVIIIGVLISVRIAVLLWSHPYIHGQDEDTN

(SEQ ID NO: 77)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q19507 (Q19507) F17A2.11 PROTEIN | 33 | 0.78 |
| Q37367 (Q37367) NADH DEHYDROGENASE SUBUNIT 1 (FRAGMENT) | 29 | 8.9 |
| PSC_DROME (P35820) POSTERIOR SEX COMBS PROTEIN | 29 | 8.9 |
| YG35_YEAST (P53273) HYPOTHETICAL 117.0 KD PROTEIN IN ASN2-P | 29 | 8.9 |
| AAF05147 (AAF05147) ORF33 | 29 | 8.9 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 452 to 1 of CT1066: this corresponds to nucleotides 241246 to 241697 of the genomic reference sequence.

WO 01/38351

86/201

PCT/US00/28888

CT1067
 Nucleotide
 Genomic coordinates:
 Start: 241774
 Stop: 243409 (SEQ ID NO: 78)

Amino Acid
 MFRQFCSLYLLQRRVNDNLRSTASASAAASLKGDGTEFITGEPPSHKMRGPSYSVLGPD
 CEDPERVYVDIVVSILQTNNIQVTKEWELFSDKLRKLGPIWIDRSGIENNGEGEEDGDENE
 DGGGNGGRIEDREAHRRKMMKKLSFVGREDPVAVDLPWTWRENSTEFARRLTLKELCDLIV
 ECGCIKSKEELFDIFEEPWEIKEAADVRGMANRSKFTKESLIDWFFEFDTYSKCVVFFE
 AVNWYLSQASPISLVLDIYCCVFSYIRRQTFLTRAKNPSLTVASSFSPTPDKLLAID
 ECVQHFLKSDINISQMALTERDCFFPLLTEMPRQKKVNTFLDTMKRPTLSLLPSTSSSS
 SSNNKRKRNTAAANILLPVYRSNFTASNNKRLKTDDGENASACILIEGYANGKISPIRI
 MVRKSTIIPVEFNHLLFPVFASKDTGANILFFIKMKSFASASLLLPGLFRHPKQFLNGPC
 KWMTLAENNINDNNINSSTMWSYTLADYCPGYYTQESPQPYQTCGNFTSTTNKRLQNVQ
 PLYF
 (SEQ ID NO: 79)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| CAB38842 (CAB38842) HYPOTHETICAL 35.2 KD PROTEIN | 39 | 0.061 |
| CAB52581 (CAB52581) CONSERVED HYPOTHETICAL TBC DOMAIN PROTEIN | 39 | 0.10 |
| O84402 (O84402) RIBONUCLEASE FAMILY | 37 | 0.31 |
| Q59085 (Q59085) RNA POLYMERASE SIGMA-54 FACTOR | 37 | 0.31 |
| O80815 (O80815) T8F5.22 PROTEIN | 35 | 1.2 |
| PI4K_DICDI (P54677) PHOSPHATIDYLINOSITOL 4-KINASE (EC 2.7.1) | 35 | 1.6 |

Comments:

TaqMan Primer/Probe Sets:
 5'start=717
 5'stop=740
 3'start=791
 3'stop=814
 5'primer=TGAAGCAGTCAACTGGTACTTGAA (residues 717 to 740 of SEQ ID NO: 78)
 Tm5=57.74
 3'primer=TTTGCGCTCTTATGTAGGAAAAGA (residues 791 to 814 of SEQ ID NO: 78)
 Tm3=58.08
 probel=ATCTCAAGCGTCTCCAAT (residues 741 to 758 of SEQ ID NO: 78)
 probelstart=741
 probelstop=758
 direction1=Forward
 Tm1=68.99
 score1=1.99
 length=98

CT1069
 Nucleotide
 Genomic coordinates:
 Start: 244241
 Stop: 244856 (SEQ ID NO: 82)

Amino Acid
 MDLSFTLSVVSAILAITAVIAVFIVIFRYHNTVTKTIETHTDNIETNMDENLRIPVTAEV
 GSGYFKMTDVSFSDTLGKIKIRNGKSDAQMKEDADLVITPVEGRALEVTVGQNLTFEG
 TFKVWNNTSRKINITGMQMPKINPSKAFVGSSNTSSFTPVSIDEDEVGTFVCGTTFGAP
 IAATAGGNLFDMYVHVITYSGTETE
 (SEQ ID NO: 83)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| O85179 (O85179) FLAGELLIN A | 34 | 0.60 |
| FLAA_CAMJE (P22251) FLAGELLIN A | 34 | 0.60 |
| BAA83944 (BAA83944) UNKNOWN | 34 | 1.0 |
| FLAB_CAMJE (P22252) FLAGELLIN B | 33 | 1.8 |
| AAC25644 (AAC25644) FLAGELLIN A | 33 | 1.8 |
| AAC25648 (AAC25648) FLAGELLIN A | 33 | 1.8 |

Comments:

EST confirmation of the predicted transcript and hits to public SBV sequences:
 Nucleotides 1 to 803 of CT1069: this corresponds to nucleotides 244132 to 244934
 of the genomic reference sequence.
 Hit to public sequence gi|6856162|gb|AF173993.1: CT nucleotides 1 to 615 match
 nucleotides 323 to 937 of the public sequence with a 100% homology, a score of
 1219 and an Evalue of 0.

WO 01/38351

89/201

PCT/US00/28888

CT500
 Nucleotide
 Genomic coordinates:
 Start: 2425
 Stop: 1537 (SEQ ID NO: 84)

Amino Acid
 MKNSRQRSGVWRGNSCLYKSFYFSGAIECKKIRIIMMFLLSLILFVCFVGVVGVIFM
 SRPNKTTTTSNKKTKKDKKEKEDDTGAVLGRREPENRPIGRDEEGAVEDGKEEEEVFE
 FEQPSVNTGSNTGGGGTGTVPGEGLPPPPPTPTPTPPPTPTPTPPPPPTPTPSPSSSLG
 EDDDDIDIDIDFDNDIDIDFLDSGEEMEEDEEEDLDTLRSLETGMSGEEVDFDASSAYI
 QPDPVVVKNIERSDYTLDPMESWKVLRSEGDIRFFVDRGITNKIKAMTEDLKL
 (SEQ ID NO: 85)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| Q48373 (Q48373) CHITINASE PRECURSOR | 64 | 2e-09 |
| O86476 (O86476) CLUMPING FACTOR B PRECURSOR | 59 | 4e-08 |
| O92451 (O92451) ACMNPV ORF91 | 58 | 9e-08 |
| Y091_NPVOP (O10341) HYPOTHETICAL 29.3 KD PROTEIN (ORF92) | 58 | 9e-08 |
| Q69023 (Q69023) (B95-8 ISOLATE) U2-IR2 DOMAIN ENCODING NUCL | 57 | 2e-07 |
| Q42421 (Q42421) CHITINASE PRECURSOR | 57 | 2e-07 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 244 to 513 of CT500: this corresponds to nucleotides 1979 to 2248 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=419
 5'stop=440
 3'start=505
 3'stop=523
 5'primer=TGCTTGGAGAAGGTTTGTACC (residues 419 to 440 of SEQ ID NO: 84)
 Tm5=59.11
 3'primer=GAGATGGGGTTCGTGTCGG (residues 505 to 523 of SEQ ID NO: 84)
 Tm3=59.88
 probel=CCTCCTACTCCTACTCCT (residues 448 to 465 of SEQ ID NO: 84)
 probelstart=448
 probelstop=465
 direction1=Forward
 Tm1=69.04
 score1=1.95
 length=105

CT501
 Nucleotide
 Genomic coordinates:
 Start: 7645
 Stop: 7042 (SEQ ID NO: 86)

Amino Acid
 MTMWNKIVITTKRMNWPVVGVFFILAITALAVLYIRHASKQEKYSTSHINEQFTAKQL
 PVTYLSKTGKLGKMDHLTHSDFMAYVDVHNRTKTLKHPMCTDEAGWAHFCLLASAEAYRRI
 RYGRGEFGPEKHS LAETIQSTVQDMSEPYITHIFKKNTDVDGHGMQSVLEKNRNKIRMGD
 GKTSSSETYNLSDKSISIVGV
 (SEQ ID NO: 87)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q21859 (Q21859) R09D1.3 PROTEIN | 34 | 1.0 |
| Q9Y3S0 (Q9Y3S0) EMDC II PROTEIN | 32 | 4.0 |
| AAD25099 (AAD25099) METALLOPROTEASE DISINTEGRIN CYSTEINE-RI | 32 | 4.0 |
| O31548 (O31548) YFJL PROTEIN (RIBOSOMAL PROTEIN L6-LIKE PRO | 31 | 5.3 |
| Q44602 (Q44602) PHOSPHORIBOSYL ANTHRANILATE TRANSFERASE | 31 | 6.9 |
| CAB52230 (CAB52230) HYPOTHETICAL 33.8 KD PROTEIN | 31 | 6.9 |

Comments:

EST confirmation of the predicted transcript:
 An isolated EST has sequence identity to nucleotides 587 to 1 of CT501: this corresponds to nucleotides 6996 to 7582 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=245
 5'stop=266
 3'start=315
 3'stop=332
 5'primer=TGGCATATGTTGATGTGCACAA (residues 245 to 266 of SEQ ID NO: 86)
 Tm5=59.44
 3'primer=AGCAGGCCAAAGTGGGCC (residues 315 to 332 of SEQ ID NO: 86)
 Tm3=59.78
 probe1=ATGTGTACTGACGAGGCT (residues 292 to 309 of SEQ ID NO: 86)
 probe1start=292
 probe1stop=309
 direction1=Reverse
 Tm1=68.96
 score1=1.96
 length=88

CT502

Nucleotide

Genomic coordinates:

Start: 8502

Stop: 7641 (SEQ ID NO: 88)

Amino Acid

MSSGSINNHPSSNMDTNKMEEGEEQDFDVLELDYSKIIHDITAMLSVAAPPPNSILDASD
 GLIATASATAPAAETGNSNRMLDKDVCQLIERDIELVKSDTIEVDSIIRQLLYFGESAS
 EKNIKTNSTEKEPVYFPKEPKGEAVKLAKNTPVLDITITKLDWMANICQSNKIGVENLASA
 LQSGQLIWTTFFAAVYASLDSFYHIAIMWKLGSFINIEALSKGSKDNLLPRDDIQVVHA
 KQEIAAMLQSRQNILGRGPSEYPPVPITAILSRITIIPLLRNFSEKL
 (SEQ ID NO: 89)

Top Blast Hits

| Sequences producing significant alignments: | Score | | E |
|---|--------|-------|---|
| | (bits) | | |
| 043631 (043631) SPINDLE POLE BODY PROTEIN SPC98 HOMOLOG | 38 | 0.063 | |
| 060853 (060853) PROTEIN ENCODED BY SACCHAROMYCES CEREVISIAE | 38 | 0.063 | |
| 060852 (060852) PROTEIN ENCODED BY SACCHAROMYCES CEREVISIAE | 38 | 0.063 | |
| AMPR HUMAN (P15514) AMPHIREGULIN PRECURSOR (AR) (COLORECTUM | 34 | 1.2 | |
| BYN_DROME (P55965) T-RELATED PROTEIN (TRP) (BRACHYENTERON P | 33 | 2.1 | |
| AAB32396 (AAB32396) T-RELATED PROTEIN | 33 | 2.1 | |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 753 of CT502: this corresponds to nucleotides 7704 to 8456 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=315

5'stop=336

3'start=392

3'stop=412

5'primer=TGACTCCATTATTCGCCAACTG (residues 315 to 336 of SEQ ID NO: 88)

Tm5=58.70

3'primer=TGGGGAAGTAACTGGCTCCT (residues 392 to 412 of SEQ ID NO: 88)

Tm3=58.70

probel=TGGAGAATCTGCATCAGA (residues 345 to 362 of SEQ ID NO: 88)

probelstart=345

probelstop=362

direction1=Forward

Tm1=68.85

score1=1.85

length=98

WO 01/38351

92/201

PCT/US00/28888

CT503
Nucleotide
Genomic coordinates:
Start: 9248
Stop: 8552 (SEQ ID NO: 90)

Amino Acid
MDPGASAASRRALWSSTVTNTRHYQQQLNRALNKIEEEDDVEEEHGQVTTTNKEMASTST
SSSSSSSSSPTSSAIPSSDEEEEEEEYDSESDTNVDSLGGEEEEEDSDTESTSADANFL
RSSRNSTTRNRLIKKYVDRFIKYEKDILLADRNKRKKRHRNRQPIHKLNNKRLKKPTD
KKQKTNNKKTWRRLPKFIKKMSPASRLKFFSACIIISGIKITSIIVLISIMAL
(SEQ ID NO: 91)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| NSR1_YEAST (P27476) NUCLEAR LOCALIZATION SEQUENCE BINDING P | 56 | 2e-07 |
| FKB4_YEAST (Q06205) FK506-BINDING PROTEIN 4 (PEPTIDYL-PROLY | 55 | 4e-07 |
| GAR2_SCHPO (P41891) GAR2 PROTEIN | 52 | 5e-06 |
| O95367 (O95367) CBF1 INTERACTING COREPRESSOR CIR | 52 | 5e-06 |
| SR40_YEAST (P32583) SUPPRESSOR PROTEIN SRP40 | 51 | 6e-06 |
| O08904 (O08904) BRAIN X-LINKED PROTEIN (BRX PROTEIN) (FRAGM | 50 | 1e-05 |

CT504

Nucleotide

Genomic coordinates:

Start: 13936

Stop: 9328 (SEQ ID NO: 92)

Amino Acid

MAHKLLFLEEEDAKEIGTLSHPEPSFALYESETFRSVGFCCKNVTDAYPKFLPRPMDINSV
 QALAVRLALIQFYKGRGWKKNMSIIDLVKOKVERNFKVDKKTSGGFIIGDGTGVGKTREL
 AAFVMSVILQEKALLDVQKHVGPSIFGQSDSKVITAINSVWKRHPFFIWLTC SKPLFNS
 CQQGMREVVTNSRGLRDPKFSWRKLQVPCANKPTSFKSDGKSGSMTVDVENS SVSSAKDSV
 DIRFFTLDVKEFHSKRSSRSIGDFTLTPTILFMTYSDLRTNLEFVLKFTTGGTDLDSN
 KVPIDNFVTTALLCDEFHKTQNISSDFRKE LAKTWEEEDTRVLRNIQKRANPSVSDLINR
 FKSAMSDDRNFKVKRMKSSNNKGRVTMSNYLKL L SQADA FRIFLEILKYDTFTVMASATP
 FQSNADLHMIDHILRKSAPAYTSIQAFKEVSSATPDAMAEHSEYVTVFLEQVIKLLRNRG
 QLVSRISISMAGVDCSTTNCKASPLQKYAIDELASYCLNARQVLIDSEKVGGHVRRRAFTKI
 IREHQEGGILEEEDVEKLVAEINSPSRKRKRAANDDDLYEVMENIDRRFKVVVVRDRDVA
 HDGKTTLR SIVQDAIKTYSQKKDALSNGGGGIITSPEVDISSIDMVAQDLYDAIKKKEKP
 SKGKTDFNEDYDDGANEEDGWGEVFDDEC FEKLRRQYFINTASTSVAACKGALLNIKATS
 VDAVKRLRTTNESKMMVMSLEQTGDSFLKNLTTRILQTIAKDESDAKYGIVDVGTFDSS
 PVANTIFSGYRLLCRAVMMASAFITISLKNKTNRRTPAHVMLVPSVPDTEPLMALAGNPI
 DSITQSIGEDSNAEITNRKLCSRITNRGLFLVKNNKTANTNKCISAFNNTKEVDVIMLG
 PKGNTGLSLHDSSNNSMYAKRYHCVLDPYNAIAFLQTIGRTHRNGQLSV PQFLIFSTDS
 PAERRFFDSL DKRIKDSKAGTYADRYSNNSIDIAA VMREQFIDQGLVLKTMGNIVQIVT
 ASMTKVHLMHEFSKMMMR TNRGVAFVEGLTLENGIFTEVIVLAMHIALVVIGAQNKITS
 SDDLGHALSFTSVLPHNQILSIVKSASQFVFSNLCLHLVHFKSDCDNLLPREKRV RDAAS
 ALIDTLNTKNNEVTSKTNKIESDAPSLTALMLPSGPRNRKMDVFSNIMAYNNNNGMDFDE
 DVPDNDEDEGCPLQEEENATTALSNFPHDYDRAIKDAHQLVTVRIVGQGEKEGVIPISE
 CLDVPELDMNTNLPVVTATNVIQSLAKENPGLLFTIHNAALAHSHREGYGGSHLLGLAKK
 LSRGFINFRTQFQNLFS PKKESKIMYDIFLSVKAIMARDDRYDGLCDMRMNSMMDASFLK
 VRKKPECVFITKLLDKNFRRHIINDEEEETRERFGGEEEEEDDDEEFEDDEEEQAEREWG
 EEEGESAYDISVINDKNNTIGHDVDIILCNRKKLTLT KENS V FVNEHIDSF MVGNLIGAE
 GSLIQICFDNCTGEFEGLPKFCLYDSSSKDKDTIP
 (SEQ ID NO: 93)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| 001940 (001940) STRAWBERRY NOTCH (SNO) | 72 | 3e-11 |
| Q9Y2G9 (Q9Y2G9) KIAA0963 PROTEIN | 62 | 3e-08 |
| O75257 (O75257) R31180_1 | 62 | 3e-08 |
| 001737 (001737) F20H11.2 PROTEIN | 62 | 4e-08 |
| O85862 (O85862) PROBABLY METHYLASE/HELICASE | 59 | 2e-07 |
| O64516 (O64516) YUP8H12R.3 PROTEIN | 53 | 1e-05 |

TaqMan Primer/Probe Sets:

5'start=2131

5'stop=2154

3'start=2210

3'stop=2233

5'primer=GGTGCAC TATTGAATATCAAGGCA (residues 2123 to 2154 of SEQ ID NO: 92)

Tm5=59.03

3'primer=CAGTTTGCTCTAGTGACATGACCA (residues 2210 to 2233 of SEQ ID NO: 92)

Tm3=58.06

probel=TTCTGTCA CCGATGCAGT (residues 2157 to 2174 of SEQ ID NO: 92)

probelstart=2157

probelstop=2174

direction1=Forward

Tm1=69.00

score1=1.99

length=103

WO 01/38351

94/201

PCT/US00/28888

CT505
 Nucleotide
 Genomic coordinates:
 Start: 16983
 Stop: 14064 (SEQ ID NO: 94)

Amino Acid
 MDERRRDPLLYPTNRSRFTAQITLFTVTFVLGCIFALVCAAMAYNVAKPMSVNFQAIHE
 LGMKSKLKAVQGANPEKTL EEYLEARGRHGVEDASNYPPHPALLDMMNLTVKGNKWNVP
 SETKERNRSR FESHDLAANRSSSLVPEHHIDRLSEATIEKSNKYLDVSGKKFRQRMVNL
 KDNIEKDDTELYDSLFGVHDIHHHSASGVSGDAPPPPPSTSEGHDEDVDILAYNTGGYCS
 NPVPLKEGQTCTSVCYTSRAVRVMTFVAGGTFITHKSGEDPKPYCWSGNVPGDHIETSP
 TTGERVVKECSVHTSIVVLTDDGGWQCRPKYPTYFGGSGGTSMTACAFNPSTHKGPPPPS
 SSTPIYYDVLKKQQIRNHTEFRNSSYISKLRQSSSLAEFKIKCNDPEFLYKNPITCFCNN
 KKDVLNNDLLSQDVTKDMKFRGMYECMENPCVMPNIDPSFVTFDVSTMKCVPGVNNPQD
 SNRHAIIGDDRTPLVGTVPAMGIFLADQSKRGDQIHQQRPKSSIDETAKKIALAQAPII
 TPLNLDATNTSKNVLFVPI PSTVLPLENIPHV IIRPSSLLHRSCLAPVLNKPSSGQHRP
 FCTAPFYIEPAANVLAGNIPQKPYEHSMLATECLRNSRMVSGSVHGGSELLFSTLLSQNK
 PSSYIRTPPGGTPAPEYNSTGDQRLEEIRDFERNFNDERRLSQTEYVIKKHARGMRTSE
 IYLKSSSWDSL MKRKEFLRHIIKKSEDTFVLKEGLLMRSYGPYAATVLARDMFDLDYLKG
 KPASKTSSTLKVSNPLQYAFPTSYSVLPEEGATDDIFSVDHNRIFDSETIPSYFDCSNVT
 PGSEKLFGTSSSSSEYRVDIDDDAWGLQSFRLDHNPKSGPVVQSDPRLAFDASNISSTPE
 GATITPLSLFKKSLVEWGHKKADVQETSWFRDGVDTSEAYRRLLVETSMAVRNSWFSWFLAW
 ENKNYYFAKNSS
 (SEQ ID NO: 95)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| CUP5_GALME (Q24998) PUPAL CUTICLE PROTEIN PCP52 PRECURSOR (| 46 | 9e-04 |
| CFHD_HUMAN (Q02985) COMPLEMENT FACTOR H-LIKE PROTEIN DOWN16 | 38 | 0.33 |
| P89911 (P89911) PROTEIN KINASE | 37 | 0.57 |
| CAB39619 (CAB39619) AIG1-LIKE PROTEIN | 37 | 0.57 |
| CAB53064 (CAB53064) DJ15D12.2 (FHR-3 (FACTOR H-RELATED PROT | 37 | 0.75 |
| Q24552 (Q24552) TF125 PROTEIN | 36 | 0.98 |

TaqMan Primer/Probe Sets:

5'start=1446
 5'stop=1466
 3'start=1511
 3'stop=1532
 5'primer=CAGGCATGCAATAATAGGCCGA (residues 1446 to 1466 of SEQ ID NO: 94)
 Tm5=59.66
 3'primer=CTTTTAGATTGGTCGCCAAGA (residues 1511 to 1532 of SEQ ID NO: 94)
 Tm3=59.21
 probel=GACAGGACACCGTTAGTG (residues 1468 to 1485 of SEQ ID NO: 94)
 probelstart=1468
 probelstop=1485
 direction1=Reverse
 Tm1=69.00
 score1=1.99
 length=87

CT506
 Nucleotide
 Genomic coordinates:
 Start: 25878
 Stop: 25197 (SEQ ID NO: 96)

Amino Acid
 MASVFEDPADLFANMDLTGKVPTRPNILFFEGLLPNSGKEIMENRLIHKGKCGAFEADTQ
 LAYFFPSNNEENTKKLNIGFQIKSNCLSFFIRDFLNDWLEIKDCGPYCTFSQYMDGDKE
 IFGNSVFGQDFTIVAMDWIDKGVTFYIFVDGSDSMENMASLWMCCKLRMNANVVKVFVD
 NASPKPFSVCKTCRWEFPGPVSVYIEGHGMGHSDDLSCDEISEFLVQ
 (SEQ ID NO: 97)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| AMY_BACCI (P08137) ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1) (1, | 35 | 0.44 |
| Q85449 (Q85449) PROTEIN 8 | 32 | 3.8 |
| Q85439 (Q85439) COAT PROTEIN | 32 | 3.8 |
| VP8_RDV (P17379) OUTER CAPSID PROTEIN P8 (STRUCTURAL PROTEI | 32 | 3.8 |
| Q85451 (Q85451) OUTER CAPSID PROTEIN | 32 | 3.8 |
| Q24284 (Q24284) PLC-GAMMA D | 31 | 6.6 |

Comments:

EST confirmation of the predicted transcript:
 An isolated EST has equence identity to nucleotides 1 to 652 of CT506: this
 corresponds to nucleotides 25216 to 25867 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=287
 5'stop=307
 3'start=351
 3'stop=373
 5'primer=ACGACTGGTTGGAGGAGATCA (residues 287 to 307 of SEQ ID NO: 96)
 Tm5=58.18
 3'primer=TGTTTCCGAAGATTTCTTTGTCC (residues 351 to 373 of SEQ ID NO: 96)
 Tm3=58.28
 probel=AGGACTGTGGACCATACT (residues 308 to 325 of SEQ ID NO: 96)
 probelstart=308
 probelstop=325
 direction1=Forward
 Tm1=68.99
 score1=1.99
 length=87

WO 01/38351

96/201

PCT/US00/28888

CT507

Nucleotide

Genomic coordinates:

Start: 29077

Stop: 28330 (SEQ ID NO: 98)

Amino Acid

MIAIANHKHDVSDALVGAHGAKINMLYGKSSLSVTEAALLMFNDTALTQFAQRGYEPSI
 PTILKAALDFSLQEEPLVAATGLDVNKA PRSWPILNCR LGYIASSNYPWAEHIISGDKE
 EIKRALEEHEKNANVRFDSDNCPVCLEDFSSSTNIIRTTRCGHCIDEKCWDRLVLSTQGE
 ITRCPVCRERTSLRPDADQVKEMLVEPIVSCKRMAVPDEQVSCKRRRIGYNRYQFLINDV
 WTDESETV

(SEQ ID NO: 99)

Top Blast Hits

Sequences producing significant alignments:

| Score (bits) | E Value |
|-----------------|------------|
|-----------------|------------|

| | | |
|---|----|-------|
| Q92T52 (Q92T52) RING-H2 FINGER PROTEIN RHA1A | 45 | 7e-04 |
| CAB51420 (CAB51420) RING-H2 FINGER PROTEIN RHA1B | 44 | 0.001 |
| Q92T51 (Q92T51) RING-H2 FINGER PROTEIN RHA1B | 44 | 0.001 |
| CAB51421 (CAB51421) RING-H2 FINGER PROTEIN RHA1A-LIKE PROTE | 44 | 0.001 |
| O17099 (O17099) F42G2.5 PROTEIN | 42 | 0.005 |
| O16682 (O16682) ZK1240.2 PROTEIN | 41 | 0.008 |

Comments:

TaqMan Primer/Probe Sets:

5'start=257

5'stop=279

3'start=308

3'stop=330

5'primer=TCAATAAAGCACCTCGTTCTTGG (residues 257 to 279 of SEQ ID NO:98)

Tm5=59.26

3'primer=CCAAGGATAATTTGAGGATGCAA (residues 308 to 330 of SEQ ID NO:98)

Tm3=58.96

probel=CAATACTGAATTGTCGCT (residues 281 to 298 of SEQ ID NO:98)

probelstart=281

probelstop=298

direction1=Reverse

Tm1=68.81

score1=1.81

length=74

CT508
Nucleotide
Genomic coordinates:
Start: 30861
Stop: 29073 (SEQ ID NO: 100)

Amino Acid
MAEAAAPRYRQVLEEVLNIEPYMSFLDVFTERELALLNDIITSRNSPPVPSSSFKKLDNK
EEFRDIYFFINNTKSDSSPICGMTFINALTTVCKTFRGLYENIHDDFLVKYSLLV
MDNGFLRRETHGKIFGTGDDSRGTGFKFTSKEQAEEREKVMRRIKKLDGVLASLKKSTSS
ARSGIVFYVEKCSSVIRFLFSRIVNITS DYVAEMKKSAPLEPFDISFGYKYFVDESPCV
TKAKRLISNGNFIIGRPFSCLETSPSSVSTDFREEMNMDARSIARLNWTNEERASAYRSV
I IKSFLSSIEEEMVEEYCETTTKTVAEMAVEFVDVFIKAETIQHFQTLYSIFDTMPKFS
AEMMDNILKNVAINAEAVGSGLCGAILLWMINSRPFEEDYNYFKICLREIMVRKKTDKLC
DNLIVKRIVSHKNVITDPHEVKGYVRLCVKVS CFMEDLEAFLTKNPWLKHTYFDEKNGT
LLCYCIINKYSHTSKLVKQEKLNILKPSAKGMSPLMVCAAISSPFTTRVGIEILTNSLA
FSFINENNENVFHAAAVATSCNFLDALAKKYKNIIYDFDRSIVNARRRAMVQRP
(SEQ ID NO: 101)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| CUL1 CAEEL (Q17389) CUL-1 PROTEIN (LIN-19 PROTEIN) | 36 | 0.99 |
| O23296 (O23296) HYPOTHETICAL 74.2 KD PROTEIN | 35 | 1.7 |
| Q57577 (Q57577) CARBON MONOXIDE DEHYDROGENASE CORRINOID/IRO | 35 | 1.7 |
| GIDA COXBU (P94613) GLUCOSE INHIBITED DIVISION PROTEIN A | 34 | 3.8 |
| Q47427 (Q47427) PLASMID P15B GENES R, S, SC, SVM1, SVMR, SV | 34 | 3.8 |
| Q39068 (Q39068) CYCLIN 2A PROTEIN | 33 | 5.0 |

Comments:

TaqMan Primer/Probe Sets:
5'start=742
5'stop=764
3'start=822
3'stop=845
5'primer=TCCAATGGCAATTTTATTATCGG (residues 742 to 764 of SEQ ID NO:100)
Tm5=58.75
3'primer=GATCTAGCGTCCATGTTTCTTCT (residues 822 to 845 of SEQ ID NO:100)
Tm3=57.51
probel=CATCCGTGTCAACTGACT (residues 797 to 814 of SEQ ID NO:100)
probelstart=797
probelstop=814
direction1=Reverse
Tm1=68.91
score1=1.91
length=104

CT509

Nucleotide

Genomic coordinates:

Start: 37245

Stop: 36048 (SEQ ID NO: 102)

Amino Acid

MAGTDIISSSSSGSSSSKKGCCIVSKKGKTIKGKNIVFKTSIKTSSSSEMMKKHKRMEI
 KDMVKKCASCKKVDYSSSTLENDALRASIESTCSALNRFPEIKYGEIGEIGDVLSAIRLMA
 GCLLAKNEKSFYKFFLRGFQFDKNGFMMLSEGMKRIEKMHTKIAKKVFGGCKAAPLKEDR
 EGKIPCQEFHKPSSYKGEYTTPLPPTPAPVKVLPLLPYKNVKNKPVFVPLAVGEAKKP
 CWVHKLFSDDPEERKRLFERHQAGRRDALMEDYGVI PNNDNEAEDTERFVSNLEYQAQM
 LELLDTANMPPPPASTPVRRGRTRIVRDYDASPVSPYSSPLHTPFDPNVNLPNGSGRMV
 DRVRDGRNRTSRRTSAVMARRINQLQHQLYSSSDSDF
 (SEQ ID NO: 103)

Top Blast Hits

Sequences producing significant alignments:

| Score (bits) | E Value |
|-----------------|------------|
| 39 | 0.073 |
| 36 | 0.37 |
| 36 | 0.48 |
| 35 | 0.83 |
| 35 | 1.1 |
| 35 | 1.1 |

O01693 (O01693) COSMID T08B2

CAB52863 (CAB52863) PUTATIVE MEMBRANE PROTEIN

O01348 (O01348) ZINC-FINGER PROTEIN KLU (KLUMPFUSS PROTEIN)

Q9ZDN6 (Q9ZDN6) VIRB10 PROTEIN (VIRB10)

Q9Y2W9 (Q9Y2W9) ENDOCRINE REGULATOR

GAR2_SCHPO (P41891) GAR2 PROTEIN

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 973 of CT509: this
 corresponds to nucleotides 36148 to 37119 of the genomic reference sequence.

CT1070
Nucleotide
Genomic coordinates:
Start: 249425
Stop: 253221 (SEQ ID NO: 104)

Amino Acid

MASSGGFFTGIDDLFKTVIQEQEKQEKNKPTQAPETEPKPGPSQAPDPVPDPVPKTPTNFC
PPPPNPLPPPPPPPKPSREERLKTSKIRLNKALSDIVEATNERVDALKENQALNTEYDK
KDNFYQVLKCSITPSVPTAIIGAHVKQVAKSSEIELAVNELDIKNKCSLVYNENESLKFF
RDHENLILQIAVQLFSRHDNTKCVGAEICVKGNEKNKFVNKLVVKKLPNAPSSSSTVLEI
RGATRNLLENFNKGENNTVNENKDI PP SERANLDTTKAEISHVFSTLHRLDTKRKLFFK
GNTFYQRKPTFDNKFRWTEVIGWTESEASKQTTKSLDKPTDDNLFVLPHSFNNLADHLRL
KEKNVLYKNSTAHPGKRNYKTQETLINPQIDSAKEYKMFVFAEIDKCLDVLLAIGKNDKY
TKSTVIQYRGKFRYYLIFCYAFYALNKAHKSRAVSPLPFNFNLFSEMYCHGPFLHSASF
LSTLTFVYQHMFPPMGTAAPSVSARLMDIDSALMKGKGKGVGRDFGSPSKTSLHTRTLV
SFLGFAEMAMGMTALLSGVEVRVSPALQQRISKSLEWCDSVIFIYFTFVLFHFRFSGAK
KVSLESALRLIMGQTHAHTNKVRAAKRCRIEAAEMEGVEEEEAGLTLSYAHLLGLPYSIQ
KALGLPVPKINPLMTASSQYNLGD FVGVEQLLKAKREFPAEGETAGFLGMFDNLVKDSI
DKYYGEGAFSDVVENVKQGMQNTPYDTSSALMTPIPKAFYEEEEKDVPQQEENSTQQRYS
LNRDVEEYLMASPMKMVFVSILDKTNQKERFMSVGDIALAVWCKRNVLKKDWNEYAIAK
GNYEWLGAKMCNHLALLADLVNFGILGDLKITNKLDNTDTFHRDSRRLPSVADQKKFIKN
TSLSDRKQLALVHSCVNVSTRTHVGRVTATSWAVDALRTYTRGDKDMFAALSSSLDMYHL
GHTNSANFVPPYFSRNYLCNEQENGLWGYTRRTSEKLAKHEELGRGRLGGLNKVGVAKTELA
AAAIAISSALDMGEVEAVMDDSSKVRKIASTCLNVNAAKVSAAREKAREASIKRLLLATN
APAAGSSRNSNRFLKDLWGFFSDPDKROKLIKGEAVSVLCPTNGFLHAAVPDFVIEYSF
ESETSIVRLRLRIKPEKQDEMVCPS TAPEANKRRKLRVNQDAVLTLDDEDNIVKYNKY
DMVEDEEARERLRHQDKQSVIAARISKVCERKNPKKKRRLEDPELQSVDEQLIRELAAIA
Y

(SEQ ID NO: 105)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| SSGP_VOLCA (P21997) SULFATED SURFACE GLYCOPROTEIN 185 (SSG | 64 | 8e-09 |
| YPRO_OWEFU (P21260) HYPOTHETICAL PROLINE-RICH PROTEIN (FRAG | 62 | 2e-08 |
| O13305 (O13305) PROTEASE 1 | 61 | 5e-08 |
| P93797 (P93797) PHEROPHORIN-S PRECURSOR | 61 | 5e-08 |
| AAD37432 (AAD37432) M-LIKE PROTEIN PRECURSOR | 59 | 2e-07 |
| Q54071 (Q54071) M PROTEIN PRECURSOR, MSZW60 | 59 | 2e-07 |

Comments:

TaqMan Primer/Probe Sets:

5' start=1884
5' stop=1904
3' start=1931
3' stop=1952
5' primer=CCGAATAGAAGCAGCGGAAAT (residues 1884 to 1904 of SEQ ID NO:104)
Tm=59.32
3' primer=TGGGCATAAGAGAGTGTCTAGGC (residues 1931 to 1952 of SEQ ID NO:104)
Tm=60.21
probe1=GAAGGTGTGGAAGAAGAA (residues 1906 to 1923 of SEQ ID NO:104)
probe1 start=1906
probe1 stop=1923
direction1=Reverse
Tm1=68.98
score1=1.98
length=69

CT1071
 Nucleotide
 Genomic coordinates:
 Start: 253296
 Stop: 255120 (SEQ ID NO: 106)

Amino Acid
 MRDDTFNQETAVKLVWRWYTEYDCCCPLVNRVERLLGSFGGGV DATSVRSRPALEYEEDKKG
 DKCIPFRITSLIEGILLERALT KPD LAAA AFDVSEKLVYCSCNNTQGNFVSSMTIWIWD
 NNSKKYEVTCPSCTVEKISGGAESIHKKPMSLLAFFNNLVEKEAFAERIELKKLYLSLLT
 GSAAGGGGMYKDSSQSSFNQSWTSLLFHTSKKDKTRLEAEVLVSNKIKHTSRLQPRCVC
 SDLLYALCSTTNNASAYAYKARNLCVIEGGEFLYFKYTI FEENGPFDSKTDLQSLVNNNEP
 VSETNSSALAASSSSLEDDDDCCDDDDDDDDDEDEKTKKKQPKKQTKKQKTTTSTLPPIS
 KTNHNDNMLMNV LKKGAVNGKRKMMDSL SGGKKQHSHKKLKTSAAGGGASSDVVAGENE
 NNPSSVSPTNNRDRKDYVLP CPQIEEVTIFSQHRMNNKLAESVVKH SVVINGNCLNLFV
 TQHRKKYILPHENILFCPLVQHVGFNKFRILTGVS CFFDRIEIVFSDQSDSVVLSNNA
 HSAILRLLSYIRENSLKR SVRTASVKGIDFVVKSDTNIGIPLSNKEIRERQLCSASTLS
 MLAGLGK
 (SEQ ID NO: 107)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q20497 (Q20497) F47A4.2 PROTEIN | 52 | 2e-05 |
| KS1_HYDAT (P38978) KS1 PROTEIN PRECURSOR (HEAD-SPECIFIC PRO | 50 | 6e-05 |
| YIL2_YEAST (P40480) HYPOTHETICAL 123.6 KD PROTEIN IN POR2-C | 48 | 2e-04 |
| FKB4_SPOFR (Q26486) 46 KD FK506-BINDING NUCLEAR PROTEIN (PE | 47 | 3e-04 |
| Q06459 (Q06459) NUCLEOLIN | 46 | 9e-04 |
| NPM_XENLA (P07222) NUCLEOPHOSMIN (NPM) (NUCLEOLAR PHOSPHOPR | 46 | 9e-04 |

Comments:

TaqMan Primer/Probe Sets:
 5'start=1007
 5'stop=1032
 3'start=1084
 3'stop=1105
 5'primer=AACTAAGAAGAAACAACCCAAGAAA (residues 1007 to 1032 of SEQ ID NO:106)
 Tm5=57.02
 3'primer=TCAACATGTTGTCGTGATTGGT (residues 1084 to 1105 of SEQ ID NO:106)
 Tm3=57.44
 probel=CACTTCCACCTATCAGCA (residues 1064 to 1081 of SEQ ID NO:106)
 probelstart=1064
 probelstop=1081
 direction1=Reverse
 Tm1=68.99
 score1=1.99
 length=99

CT1072
 Nucleotide
 Genomic coordinates:
 Start: 255074
 Stop: 257477 (SEQ ID NO: 108)

Amino Acid

MLSFNPEYASWFGKMITDPGVILPVSKDVVLFSGRQSDVGIMTLDPHLDIKITSKRIG
 VEERLAQYNTLPMDFTRAMEKELNSRNMKESIFTGIFLDTGSAIFEDNMFNGGGSALRL
 IRSPALNSAVFSSKNYIIKQLPTITKSLRRSQARDKQVDKREKIVVDSFSILSAIAAQV
 MHLTDGEMTYVPDGHCVNVVMSETNASSIYLIINDPTGSGWKIMPNNFNKTLEMRDGVID
 RVETLVEFACKCVASSLIKRGMDLVDQMORTIRSMDFLPPASSTSNNTPRVAIMTSGSSTT
 TGIGSLSILAEDGSTHHQIKLSEYRTGLSITENNREVSFTVEPSIDGVQAEHPLSPSILQ
 WLPPLVKRPEVVAAAAA AVVEEENGDNKPSDKDNEDKYSDFWSNVPVTPLITPKKWRA
 CKINDRAMISSWKNNLVKLHKYDWTNKTTKVDYFDKMAAFVALMTFRKFQDILADNYVPP
 QTPSQGSEYAVTMSNVATLFTDVYGFESNGNKPLFALEQLENETGIESIYVLNIIGNSPD
 GNSVRVVRLEKEMSELLKAKQYFTEMAIPPINEKCKWTDKAPSSVKEYKYFCDLTAPISK
 RPRKDNNDGGVEHSALTYTPRCIYHTERCLVHLYSEPEKITEHVSFNKDLNILEIGKNIT
 NOYQNTYKSIFEIVDVPIIVASMSSTKTMTVNNYIIISTPSATTKFVQDPPKTGKQLLAVE
 EVRNFKLKSVLVPPPYFRDNKRNTTLCSEQITEQNCPSSSEGGRFSCPSSESLILKYSNLSK
 KRALEEIAPETETSILSLAM
 (SEQ ID NO: 109)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| TFC3_YEAST (P34111) TRANSCRIPTION FACTOR TAU 138 KD SUBUNIT | 39 | 0.12 |
| Q9XIC1 (Q9XIC1) F13F21.2 PROTEIN | 38 | 0.20 |
| P95629 (P95629) PUTA GENE | 37 | 0.60 |
| Q9XGN1 (Q9XGN1) TTG1 PROTEIN | 35 | 2.3 |
| O85475 (O85475) CELL DIVISION PROTEIN | 34 | 3.0 |
| Q00741 (Q00741) TAMA | 34 | 4.0 |

Comments:

EST confirmation of the predicted transcript:
 An isolated EST has equence identity to nucleotides 1 to 932 of CT1072: this
 corresponds to nucleotides 256520 to 257451 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=1048
 5'stop=1068
 3'start=1099
 3'stop=1120
 5'primer=GCAGAGCATCCTCTATCCCCT (residues 1048 to1068 of SEQ ID NO:108)
 Tm5=58.38
 3'primer=CTGCTACCACTTCTGGCCTTTT (residues 1099 to 1120 of SEQ ID NO:108)
 Tm3=58.57
 probel=TTCAGTGGTTACCTCCTC (residues 1076 to 1093 of SEQ ID NO:108)
 probelstart=1076
 probelStop=1093
 direction1=Forward
 Tm1=69.00
 score1=1.99
 length=73

CT1073
 Nucleotide
 Genomic coordinates:
 Start: 257551
 Stop: 259132 (SEQ ID NO: 110)

Amino Acid
 MGVOKNILVGGGGVSLLLGVVTLTGTVTEGAPAVPPFSSSSYSFTPESSVFWVEGNRVL
 SGTKKDTLINVLGKKIPYYANSIFRHDCSETRSIQWPETSPLGLNLIFCSCASHEHQHRT
 HETTEPDDLLWDGSRKTTTIIIPKKWSDVWVWTSWLRDNDQKCGCGQAFVSSFTSTQKEV
 QGEWLAHTNGKTSEGDNSAYLFIQLRTTLKPIITDVTEDNMMMGMSGTPMNPKDMT
 YFVNDFSDDIGSTPQCLVNSDILNKREEWIAVWGVADSKDLLTKHQLGEREYGSEGRRR
 NPGVEEEEEERVEEEEEVEALPYIKKSGKLIGPRRRPLTTTTTTTTTTTTNPIVREVVE
 DFDYESFNEPEIFGSNSKLPFIRFLDQKNWRLGIMSRVSSIANFKIEQESSKALFCLAV
 WVGDEHTPKFRLSVWKNWKPFTSAPIIVQNVGYSSDVFWHETLRSKIVDRSRDLIETKVT
 KKIGEDWANKKQTVVAMFISGIVCITVTVISIFSIVIYYIKIMPKF
 (SEQ ID NO: 111)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|--|-----------------|------------|
| G1184543 (G1184543) NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, | 44 | 0.003 |
| O17412 (O17412) CHITINASE | 41 | 0.015 |
| AAF00095 (AAF00095) HISTONE ACETYLTRANSFERASE MORF | 39 | 0.10 |
| AAF00100 (AAF00100) HISTONE ACETYLTRANSFERASE MORF BETA | 39 | 0.10 |
| O15087 (O15087) KIAA0383 (FRAGMENT) | 39 | 0.10 |
| AAF00099 (AAF00099) HISTONE ACETYLTRANSFERASE MORF ALPHA | 39 | 0.10 |

Comments:

EST confirmation of the predicted transcript:
 An isolated EST has equence identity to nucleotides 511 to 925 of CT1073: this corresponds to nucleotides 258607 to 259021 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=698
 5'stop=721
 3'start=823
 3'stop=846
 5'primer=CCATGAACCCTAAGGATATGACAT (residues 698 to 721 of SEQ ID NO:110)
 Tm5=57.04
 3'primer=GAGGTCTTTAGAGTCTGCAACACC (residues 823 to 846 of SEQ ID NO:110)
 Tm3=57.29
 probel=TTCCGACATCCTGAACAA (residues 780 to 797 of SEQ ID NO:110)
 probelstart=780
 probelStop=797
 direction1=Forward
 Tm1=69.00
 score1=1.99
 length=149

CT1074

Nucleotide

Genomic coordinates:

Start: 274526

Stop: 275153(SEQ ID NO: 112)

Amino Acid

MYIFVEGSPLTGKSSWMSKLIDTGSCGMSFLNFLRMNTSDYYNWP AEIGTEHLQLGFRET
 RVVDGMFEPVLKTFVDSWKKEQGKESLKEYLDYNGQVMEIYIAEWLRQRPLAFHVFTYTD
 EAVKSGFLNEEDLDMDTATKWMAEIIREKRGNIQEIKVTPRVVFNGNVCSACFSNTKRNL
 YNFGTNYNNVVHCDLLCPFARHRIVHFL
 (SEQ ID NO: 113)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q9Y7X6 (Q9Y7X6) HYPOTHETICAL 75.7 KD PROTEIN | 38 | 0.079 |
| Q9YBC0 (Q9YBC0) 431AA LONG HYPOTHETICAL PYRUVATE DEHYDROGEN | 35 | 0.52 |
| O67329 (O67329) DIHYDROOROTATE DEHYDROGENASE ELECTRON TRANS | 33 | 1.5 |
| DPO1_BORBU (O51498) DNA POLYMERASE I (EC 2.7.7.7) (POL I) | 32 | 4.5 |
| YHC3_YEAST (P38742) HYPOTHETICAL 130.0 KD PROTEIN IN SNF6-S | 32 | 4.5 |
| O51342 (O51342) ATP-DEPENDENT CLP PROTEASE, SUBUNIT A (CLPA | 31 | 5.9 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 563 of CT1074: this corresponds to nucleotides 274555 to 275117 of the genomic reference sequence.

WO 01/38351

104/201

PCT/US00/28888

CT1075

Nucleotide

Genomic coordinates:

Start: 277704

Stop: 278079 (SEQ ID NO: 114)

Amino Acid

MWRSCISNIREMGDNKDYETRLIQRINDLESEIENKTELCEKINEQMKNLQKYDKCFVE

EETEKFRKMEERVLYLKEQGIPLOPEERRTMLAEIDKSNKELDALLEENERI IKLIDEEL

ESMK

(SEQ ID NO: 115)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| O44929 (O44929) MICROTUBULE BINDING PROTEIN D-CLIP-190 | 47 | 5e-05 |
| Q13439 (Q13439) TRANS-GOLGI P230 (256 KD GOLGIN) (GOLGIN-24 | 43 | 7e-04 |
| Q59037 (Q59037) HYPOTHETICAL PROTEIN MJ1643 | 43 | 0.001 |
| Q25662 (Q25662) REPEAT ORGANELLAR PROTEIN | 43 | 0.001 |
| KINH_DROME (P17210) KINESIN HEAVY CHAIN | 42 | 0.002 |
| AAD29948 (AAD29948) MYOSIN HEAVY CHAIN | 41 | 0.002 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 388 to 1 of CT1075: this corresponds to nucleotides 277689 to 278076 of the genomic reference sequence.

WO 01/38351

105/201

PCT/US00/28888

CT1076
 Nucleotide
 Genomic coordinates:
 Start: 278221
 Stop: 278473 (SEQ ID NO: 116)

Amino Acid
 MQKKYDKLFEDDKRFREIEERILQQKEKGNPLDPEERLVLSADIDRSMKEIDDCLEEINH
 IELSIDTLTLLDECENLHYGLQTTK
 (SEQ ID NO: 117)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| O18164 (O18164) Y6B3B.1 PROTEIN | 35 | 0.076 |
| P70388 (P70388) RAD50 HOMOLOG (S. CEREVISIAE) | 35 | 0.10 |
| Q06268 (Q06268) INTERMEDIATE FILAMENT PROTEIN | 33 | 0.30 |
| USO1_YEAST (P25386) INTRACELLULAR PROTEIN TRANSPORT PROTEIN | 33 | 0.39 |
| Q07380 (Q07380) HYPOTHETICAL 206.5 KD PROTEIN YDL058W | 33 | 0.39 |
| O96275 (O96275) RESA-H3 ANTIGEN | 32 | 0.51 |

Comments:

Hit to public SBV sequence:

gi|6007410|gb|AF178573.1: CT nucleotides 13 to 252 match nucleotides 2050 to 2289 of the public sequence with a 100% homology, a score of 476 and an Evalue of 1e-137

TaqMan Primer/Probe Sets:

5'start=88
 5'stop=111
 3'start=147
 3'stop=170
 5'primer=AACCCTCTAGACCCAGAAGAAAGA (residues 88 to 111 of SEQ ID NO:116)
 Tm5=57.69
 3'primer=TCCTCGAGACAATCATCAATCTCT (residues 147 to 170 of SEQ ID NO:116)
 Tm3=57.97
 probel=TATTGTCGGCTGATATTG (residues 116 to 133 of SEQ ID NO:116)
 probelstart=116
 probelstop=133
 direction1=Forward
 Tm1=68.98
 score1=1.88
 length=83

CT1077
 Nucleotide
 Genomic coordinates:
 Start: 278723
 Stop: 280976 (SEQ ID NO: 118)

Amino Acid
 MEKKTETAATTEKDPEPSVSKRSRNKEPKTTSTVYTSVKCYLSSIIKSESSRSNVTSTKE
 RFEERCKSVSKMMVKGSLFLRLVVDECLRRYNHLEDEIDKWPDMTKDNFYVQLLRKGLDK
 KKLKEGSTHPVVEDVWNSPIVQETFLSQQEGGNPIKRHLMDFNITITYAAKQLKTCFETN
 LRTHFRTRQQRASGWLAEANGFDKKYTKLVQHWIIGCTYKSDWVDSGDLERVKEGTKNFV
 TLHRKHLCVISDKKNGTISYSPEEKYPIPSILNYYKFLQTEYPQNKKIQKMIVVPKHKLK
 IHYCTFDQTTIQGICKDLGVWKDMEERHKQSEDI LYKQGWYLLFDVKKIKKL RPNWNFHS
 IQTDGEGSVLFSREVEEVETVSKKSKKNKKPRGDEDRNYPPTNAKYVVGVDPGRTNVV
 SCSVFDTRQKRVRKHRMTAKQYYQESWMTDRRKANET YKKNKEYKEALEEITRYDNGE
 EIINDGNGDTSTPTKKFEAYLKVVNEHYRLLWNEKGKKKYRK NAMKVYSRKQK CISNFI
 ELIPKRDKIEDYHIAFGDAKFACTGRGEQYASPARIFAKKIKERVGGDKRFTFVDEKYTS
 KVCHRCNQPLNMLEKDCFS PNKKRPPTIVTTTTTTTTTEEDEENGKWKATPLRENDRTR
 RCSSEKTQFGYSSNRKVSTGDISMETPVPSSSTSSSFCTPT SITCVLGGKFVDRDFNASTN
 IVHKFLGFWDDKKLMEKKDKMPLKYHFIRVA
 (SEQ ID NO: 119)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| Q9ZAY5 (Q9ZAY5) SURFACE PROTEIN C | 46 | 7e-04 |
| CYL1_HUMAN (P35663) CYLICIN I (MULTIPLE-BAND POLYPEPTIDE I) | 41 | 0.022 |
| O62231 (O62231) F35E2.9 PROTEIN | 39 | 0.086 |
| Q12080 (Q12080) P2610 | 39 | 0.11 |
| MYS2_DICDI (P08799) MYOSIN II HEAVY CHAIN, NON MUSCLE | 39 | 0.15 |
| O69188 (O69188) C3-BINDING PROTEIN | 38 | 0.33 |

Comments:

Hit to public SBV sequence:

gi|60074101gb|AF178573.1: CT nucleotides 12 to 297 match nucleotides 2549 to 2833 of the public sequence with a 99% homology, a score of 551 and an Evalue of 1e-159

WO 01/38351

107/201

PCT/US00/28888

CT1078
 Nucleotide
 Genomic coordinates:
 Start: 282175
 Stop: 282586 (SEQ ID NO: 120)

Amino Acid
 MGNSESRSSGIEIVHKNGAPKRSHKTYLSNRTERHAQIQKQIEELHHKTNKQFEQAQKV
 LDKNEERKKHQQQQIIIPLDPEERRAILAEIDKHMKEIDGFIEESERLGLLVDAEINNLL
 EEKEVEEEHLLKQKED
 (SEQ ID NO: 121)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| ARP4_STRPY (P13050) IGA RECEPTOR PRECURSOR | 45 | 2e-04 |
| AAF05247 (AAF05247) ORF133 | 43 | 6e-04 |
| MYSG_CHICK (P10587) MYOSIN HEAVY CHAIN, GIZZARD SMOOTH MUSC | 43 | 8e-04 |
| AAD46501 (AAD46501) LATENT NUCLEAR ANTIGEN | 43 | 0.001 |
| O40947 (O40947) ORF 73 | 43 | 0.001 |
| Q9ZGM5 (Q9ZGM5) M-LIKE PROTEIN (FRAGMENT) | 42 | 0.001 |

Comments:

EST confirmation of the predicted transcript:
 An isolated EST has equence identity to nucleotides 5 to 344 of CT1078: this
 corresponds to nucleotides 282190 to 282529 of the genomic reference sequence..

CT1079
Nucleotide
Genomic coordinates:
Start: 286863
Stop: 289635 (SEQ ID NO: 122)

Amino Acid
MSSSSSSSFSFRISTYQTFLLKALAHPLVDKITQKCDTGRNQKCPFIQFLADISHLIQGE
RNGGNLFLPLHPFNQPHLEPRIVGSLHGRTLDNDIEESYCYFVKDLYNGVFSYVNGVKEL
QGVLDKKISGSGSGESSSSRAPLIPITDVLVLLYIFGTLVLPVPPRSKAYRVITEAVLALPF
NEFSNNWPPTNIKAYVSRDFRMFNLLAGLDHIEGEVGGSEWESI HASVVKRMVTIMRN
KAEKKPPSTSRIFRVYVAEPVNDVTKIPIRVLSKLFGRLAGILQKVYSYMLNLPYLL
SSNSIDIKQGVKGITLSIPSARKLGFYLLQKDTTLQSSLSQDVADCIVSINAGIIGDDFS
EKIROCIEEKNKPCNCCMCFCEIDKTPDFSYSEHVARHNFFPVHAFSSSHDDKCCGAKIC
SECIFPYIISLYEKMGTGAVGVKVVDFLQCPGCKSGMLNLKGRCYEFSNLCKRMILPYTST
HCSSLFDATINRAEACFYSLFLQYDFETARRIAHGAKDIPHYNVKVVKNVKDLRLCAL
YCYKCVSPVVCDEPNESTDYEMVDVTPPLINLIEVDSEYDDGPGNHMWPAKFTCNFIA
GSSGETPTISTCRDAVTFGLRAPRKKMAGWDDQSAVGQAI ALANWRKSGELPKNMFDLL
EGVNAVLYRGDSFLLRAINYPVIGRSMSPSLELVKRVKNKIALIAKFAFFHEKRVRPDASK
KLEWAELLVKSYLEVLLQTPECVIHRAHSFVGKTLITDELVHMRPDDATRNAYIQNL
NAARONAAAAASFSGSLPKPEFVPCKERTIEWMYEKDNDVVRVNCPSCKKAIQKYGGCV
NVFCECGTNMCWICEEKVSPADSNHCVEKHRIVYSNCVRVKYALESMYGFECTMKNVEE
GVKNYYVMENGFFFDVQEMVAKK
(SEQ ID NO: 123)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| AAF04637 (AAF04637) HYPOTHETICAL 84.4 KD PROTEIN | 213 | 3e-54 |
| O97226 (O97226) PFC0175W PROTEIN | 49 | 1e-04 |
| Q94981 (Q94981) ARI PROTEIN | 48 | 4e-04 |
| CAB45785 (CAB45785) HYPOTHETICAL 262.6 KD PROTEIN | 46 | 0.001 |
| CAB36714 (CAB36714) HYPOTHETICAL 68.5 KD PROTEIN | 45 | 0.002 |
| Q9XII0 (Q9XII0) F7H1.11 PROTEIN | 45 | 0.003 |

Comments:

EST confirmation of the predicted transcript:
An isolated EST has equence identity to nucleotides 1 to 910 of CT1079: this corresponds to nucleotides 288641 to 289550 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=1148
5'stop=1167
3'start=1192
3'stop=1211
5'primer=TTGACAAGACGCCCGATTTT (residues 1148 to 1167 of SEQ ID NO:122)
Tm5=59.14
3'primer=TGGACGGGGAAGAAATTGTG (residues 1192 to 1211 of SEQ ID NO:122)
Tm3=59.35
probel=AGTGAACATGTGGCAAGG (residues 1174 to 1191 of SEQ ID NO:122)
probelstart=1174
probelstop=1191
direction1=Forward
Tm1=69.00
score1=1.99
length=64

CT510
 Nucleotide
 Genomic coordinates:
 Start: 38917
 Stop: 37381 (SEQ ID NO: 124)

Amino Acid
 MAETVAVDEVPTCPICMGDYDSDDTCYNWSNGGMPCCRKSVHLECLFTWRFEHVMVNEH
 LLCPCMCRAYIPVWFFRKVYEEVYKYASFHSFLLSADYVNDEGVKDTLNKMSTILAPTFF
 VPNAKGVNENEDVYMERAYTKLSFMLETLSRQEMHAFSEETFEDNHEAALMGKFKDIPPY
 EYEGEWLKYVAPNTIDITQCLSNDDDDDEGDNNVSPSLLSGVTSFNFIEDDEDTVVFVPP
 EVDDNDDSESLPDLTVPPRSNNITFDTISGISSSLYDVNDDDDDDTMSLPDLNMPASST
 SSAPTSSAPTSTSLNINVNLCFNVDSDDDEEVIPSSSSVNQPSTSSGSSSSSSNSRKR
 RYGRDEDMSNISSESKRLCVDVKRYMCRLDNIDEEYNEIANRYLAELSALRERRQETEN
 KLGDCISRGNLFHTTVNDVIGKSLCSKKLKVKKRYASKWSANKQLIGSCLIKSASNNARL
 DDEIAHVHSSLLNGFDTDPSEADQISSLPNL
 (SEQ ID NO: 125)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|--|-----------------|------------|
| O70567 (O70567) DENTIN SIALOPHOSPHOPROTEIN PRECURSOR | 56 | 7e-07 |
| P97399 (P97399) DENTIN SIALOPHOSPHOPROTEIN PRECURSOR | 56 | 7e-07 |
| SR40_YEAST (P32583) SUPPRESSOR PROTEIN SRP40 | 53 | 4e-06 |
| O95815 (O95815) DENTIN PHOSPHORYN (FRAGMENT) | 53 | 6e-06 |
| P87736 (P87736) RING-FINGER PROTEIN (FRAGMENT) | 51 | 2e-05 |
| Q53653 (Q53653) CLUMPING FACTOR | 51 | 2e-05 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 444 of CT510: this corresponds to nucleotides 37411 to 37854 of the genomic reference sequence.

CT512

Nucleotide

Genomic coordinates:

Start: 61872

Stop: 60672 (SEQ ID NO: 128)

Amino Acid

MLSTCDLKHPSSTDGNVLKNIHFSESIPANDIISFPSSDTEELNKDLLDSVRNQIKFGFD
 PITETLKNCITTQTLHLSFLKSSLLTLQEKFNWGSIQLEKGGQEMALCASLKIMQISA
 LIETAKEASMDNKKNNNACANCRDSKCSASLVTLFNKTIDEKYVKQNSSSASALLANTF
 TAGANKPPKEFITKDNAHGNSDNTYTAMSDNLICPGKYSSDITYEVTQAKERIKNNNK
 KMRLATGVEMVMKELEAENNKEGGRVEVEVEGVEQQPSTSGEEMQMEIMLPPTPPPDLE
 SLVTEGVDDYPVFSPLPSLLSPMPASPLPSNGNSALEDGGPFAPSADIVVDKTSEIMGR
 PGSEWVHQDRNSKMEIRNYGARGSGINTGRYRRNNTVL
 (SEQ ID NO: 129)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| GARP_PLAFF (P13816) GLUTAMIC ACID-RICH PROTEIN PRECURSOR | 38 | 0.12 |
| YMEI_YEAST (Q03433) HYPOTHETICAL 32.0 KD PROTEIN IN CAT2-AM | 36 | 0.36 |
| Q9WTU0 (Q9WTU0) PHD-FINGER PROTEIN | 36 | 0.36 |
| O15029 (O15029) KIAA0312 (UPSTREAM REGULATORY ELEMENT BINDI | 36 | 0.36 |
| MMPL3_MYCLE (O06081) PUTATIVE MEMBRANE PROTEIN MMPL3 | 36 | 0.47 |
| Q06166 (Q06166) MATURE PARASITE-INFECTED ERYTHROCYTE SURFAC | 36 | 0.61 |

Comments:

TaqMan Primer/Probe Sets:

5'start=549

5'stop=570

3'start=626

3'stop=649

5'primer=TGCAAATAAACCAACCCAAAGAG (residues 549 to 570 of SEQ ID NO:128)

Tm5=57.79

3'primer=TGCCTGGACAAATAAGGTTATCAC (residues 626 to 649 of SEQ ID NO:128)

Tm3=58.13

probel=GCACATGGCAATTCTGAT (residues 589 to 606 of SEQ ID NO:128)

probelstart=589

probelstop=606

direction1=Reverse

Tm1=68.99

score1=1.99

length=101

CT513
 Nucleotide
 Genomic coordinates:
 Start: 77506
 Stop: 76273 (SEQ ID NO: 130)

Amino Acid
 MEEHLSFNKSPENGVVFFDFSDNTSMSNMVDNIRHRLPMDKKFSSKALLLASTPIPSDE
 QLSTKVNKAIIFSHRETIVLSKALKIVVTGLYVDGEYVDDVICLYPEKHTLNGILRYVVHL
 NMMLMDKAEDADEIRCGLIPLGRGFNREAFKVPDPIPCAGYNILNGYHPDNGHQISPSS
 TQPQVQRRCAVKQMYKQINGMFEVVKQFSIKHNNRIFTINQVDFKGEEMKMFFALYSEEL
 LPFYSETGKLLSEKHVSKSFSQLPPHVTISVFYLRNMEEYNTLMKTDFGSCFAPAIDT
 GDNFELFGMNNNIVSKVCGDDALDLRRRIMEHISDAIGRNVELADNRLNPHITHGKIN
 EGVVGEWVSRAFPCNFLCKPREIIVFGGTFIFGRVSNNGNYVIKQPVQDVV
 (SEQ ID NO: 131)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| O32743 (O32743) RECOMBINASE | 37 | 0.23 |
| SYA RHILV (P24075) ALANYL-TRNA SYNTHETASE (EC 6.1.1.7) (ALA | 34 | 1.5 |
| O03942 (O03942) LACTOBACILLUS BACTERIOPHAGE PHIG1E COMPLETE | 34 | 2.6 |
| Q9X257 (Q9X257) CONSERVED HYPOTHETICAL PROTEIN | 32 | 5.8 |
| O22993 (O22993) CELL DIVISION PROTEIN ISOLOG | 32 | 9.9 |
| YK05_MYCTU (Q10851) HYPOTHETICAL 30.9 KD PROTEIN RV2005C | 32 | 9.9 |

Comments:

EST confirmation of the predicted transcript:
 An isolated EST has equence identity to nucleotides 1 to 982 of CT513: this
 corresponds to nucleotides 76314 to 77295 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=502
 5'stop=522
 3'start=559
 3'stop=578
 5'primer=TACCACCCAGATAATGGCCAC (residues 502 to 522 of SEQ ID NO: 130)
 Tm5=58.25
 3'primer=TGCTTGACTGCGCATCTTCT (residues 559 to 578 of SEQ ID NO: 130)
 Tm3=58.35
 probel=ATCTACTCAACCACAGGT (residues 537 to 554 of SEQ ID NO: 130)
 probelstart=537
 probelstop=554
 direction1=Reverse
 Tm1=69.03
 score1=1.96
 length=77

WO 01/38351

113/201

PCT/US00/28888

CT514
Nucleotide
Genomic coordinates:
Start: 102885
Stop: 100046 (SEQ ID NO: 132)

Amino Acid
MDSTSTTTIEAEKALLKEYVNEENLTWEFVDRVIRHEKLMQRTDMRLKTSSRRLFSFISIIY
SFLQDFFTARDGVNSDEWCTQSALYHMLDGVASIIISCFRKRIDYNNKKMERLACTSIREG
YFLVDVKTIESRHVELLDPPDKKIWRQRYAEKIAPEKVVDAYNEVSKLLPDEAMANYNYRT
GLVHLSDTLKNAKKPPTDLTMTDFDFYEKYIRSDIVLGKSNKLSGMFSENFELPDINIK
VPRRLERYFNVETNYSLEHNFRFPPSNHIRGLIFAYFIGNIFGGAFSCVQLYLLGFTLSAA
SACRENVLDTPFSKLGKQYIKNDNKTKNSSSNEDNDEEYYPCELQYARINSNDKNACRKS
IVKAVKFVADRVEKASVTMMRTPIAEHESDGYMADWLSLQISKLLGRKVSASYALLFIVN
WVAHKYKQSFTNDVNGSEKYEILLKKLTVACGLTYNHKCGMVVPVIGFGSGMTNRKLRQY
AVHCIENVIGSFISGKRKKDIHEDPKKLEEMSLMQLSARLFKNNDVMKRGQDGKVTFAN
EDNVQDFLEELKTKEFVPNERRRKIHEEYTKSLHTNLKMTFRFGVCGFQHPLPASSDKP
TQVSLQLLKQRQTFVQRETA AVNWTRLLOQLFSPDERDNKRHQNLSLWNRLGSLNLRHF
ISLASKFIKRSVHCERVVNDIISKFNADILPLGKDPDHFLMTKAGLVIEDHARENIDNAM
YSLCGGFNNQTTQKLNISIRLISAEALKNARNCVLATTFSKSYNEDRPFPLRTDEAKFV
PIPLFGVEPLHPLLSFIDNTANKCNSVSDFWLEESDDIFKEALVSHITLTDSSVYSTL
VGEDEYCDNNKSKRIGNTLVCTLYDMMGRANYNGLHSDKPRKHDPWPSSKNTGQSGR
STTDFSPNSVIVLLDTENVADDYEDDEEDYEALKQSERDNVITLNNX
(SEQ ID NO: 133)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|--|-----------------|------------|
| Q12537 (Q12537) GLUCOAMYLASE PRECURSOR (EC 3.2.1.3) (GLUCAN | 39 | 0.11 |
| AMYG_ASPAK (P23176) GLUCOAMYLASE I PRECURSOR (EC 3.2.1.3) (| 38 | 0.32 |
| AMYG_ASPSH (P22832) GLUCOAMYLASE PRECURSOR (EC 3.2.1.3) (GL | 38 | 0.42 |
| MSH6_YEAST (Q03834) MUTS PROTEIN HOMOLOG 6 | 36 | 1.2 |
| YMM1_CAEEEL (P34489) HYPOTHETICAL 81.8 KD PROTEIN K01B6.1 IN | 36 | 1.6 |
| AMYG_ASPNG (P04064) GLUCOAMYLASE G1 AND G2 PRECURSOR (EC 3. | 35 | 2.1 |

Comments:

EST confirmation of the predicted transcript:
An isolated EST has equence identity to nucleotides 3 to 841 of CT514: this
corresponds to nucleotides 99939 to 100777 of the genomic reference sequence.

TaqMan Proe Sets:

5'start=1126
5'stop=1147
3'start=1173
3'stop=1194
5'primer=TCCGTCACAATGATGAGGACAC (residues 1126 to 1147 of SEQ ID NO:132)
Tm5=59.41
3'primer=GGAGAGCCAGTCTGCCATGTAG (residues 1173 to 1194 of SEQ ID NO:132)
Tm3=59.86
probel=CGAACACGAATCAGATGG (residues 1155 to 1172 of SEQ ID NO:132)
probelstart=1155
probelstop=1172
direction1=Forward
Tm1=69.00
score1=1.99
length=69

CT516
Nucleotide
Genomic coordinates:
Start: 118885
Stop: 115402 (SEQ ID NO: 136)

Amino Acid
MKIVQNNFTPDERSGVIHIRKPAKIEKAVFGNIAAAIDDSAAVRKDPKKRNLKNGLEPA
SKKLAKNIERISSEELKRVTDVQDPKLLHSIMKRTARQIGYDIGDDISQSPAPDRDGS
SSLLPIRMNIRTEELLEKGGKDTIVRIHILDGILPDNVPLPFKAEIKVDLVDEKEYEGED
GGGSSDSGSLFETFPFVPAWPPITNDPNAFSRNNGNKQQAVFKHVEVNSLADGITLS
TKGSIFNTGNRLKISIVTEDKNKTVLFDSQVTISSPIPKITEVFACRNVALMRLDMPKAI
NYDNVEYTPDITLNEKYVSDYPANFPRLSRQAEIASNLAAKLPRENQLSDINKPSVSFVYS
KTNTVNTPVNLKVLNETLKNMEGNESEGYKILNATEITHLRNPSNPARTFICVSVPESE
IEAQWKMLGWIVGFKTSSDVLTTSSSGYNIVFPASKVTQSDKLFVISTDVNANTNKVVVH
NTPSRVGCFCGSSVNFVRVDAATAPDWPGPTNGPDFFSYQLRPCIILKTDNDNREPRTAVL
SSPATEYAGERTTSLPRALNVSVGPLETEVRGGDIITPVQTALLGGEQPTFKAPAEPTKL
YAVFPVLDSHNLVKASDNPFQPIHSITSRNKTTVLTVDIVNDDDDVLEDKSYHIT
VSDPVSGSILAKENVLSSRITSRPIFDGARDDRVFSVKMEVFGGDDKGIQMPFTMDGHF
EGQFSDMSVPSNELAIWNDPSTFTAPVRDTPATDITNKGIVYCRITLPPISNRGIRDPFM
KQTSLVPLPTSIPEWAFADYGGGEIKYPRHIFISSIRTNDTNIIVNTDTQTEFSIENWLRE
QIDKEQERHRQLLPAPSEAYTQGEKVYAKMYMGDGVSEETLDQIVHTSNTTYVVDSEGTK
KENLLVNKEDKKLAAILGKWGIVVFGANKYPDEPADRYTNWRNTGRLRAVGSYSQLRQPV
APLQTRLATWPSGDPVTRLADGQFLVRLDPRCGGIGSANGFYNNNGANNEFTSSLLFAIV
GNQDKVVSYAERVRFYMKIVARNEGKKHLKNDGDLVLVDRNSALHRLWNRTTFDHDHDI
LCVKIPQNVMSKIEPGTSSGVLDPLVFANVASSTDREEFYKKFIDTSSGPVVIDRASVT
SSYNISVPLNFYTTTCGFIVG
(SEQ ID NO: 137)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| DNA2_YEAST (P38859) DNA REPLICATION HELICASE DNA2 | 37 | 0.68 |
| O94534 (O94534) PUTATIVE SPINDLE POLE BODY-ASSOCIATING PROT | 37 | 0.89 |
| BAA84527 (BAA84527) ALP14 | 37 | 0.89 |
| YGS4_YEAST (P46947) HYPOTHETICAL 30.5 KD PROTEIN IN SAE2-KE | 36 | 1.5 |
| O28907 (O28907) GTP-BINDING PROTEIN | 36 | 2.0 |
| Q49547 (Q49547) LMP3 PROTEIN | 35 | 2.6 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 888 of CT516: this corresponds to nucleotides 115494 to 116381 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=1869
5'stop=1892
3'start=1984
3'stop=2001
5'primer=CATTACAGTATCACGTCAAGGAA (residues 1869 to 1892 of SEQ ID NO:136)
Tm5=57.80
3'primer=ACCAGACACGGGATCGGA (residues 1984 to 2001 of SEQ ID NO:136)
Tm3=58.39
probel=AGCTACCACATCACAGTC (residues 1966 to 1983 of SEQ ID NO:136)
probelstart=1966
probelstop=1983
direction1=Forward
Tm1=69.00
score1=1.99
length=133

CT518
Nucleotide
Genomic coordinates:
Start: 190743
Stop: 188172 (SEQ ID NO: 140)

Amino Acid
MTESKDYVLALVAETKTDEKRLNYVSEGLVAAISNLQNTPEKQKRVV ISSDVFGPTWFNK
TTEFFNSGLRLAKGHLKDAVMRSVYRDIEGVREHIIDPSWRLTETAAEELCDFTLKQA
PLNLLNAPENIMDGVFRSAANLVLYSTRGDTNEPSWVIDSEMLANRNNSTVADLAMGRA
KRAIALFLGYTLCDILRWQSIASRMKERGLDPFAAMPPHLEYGRAADMIEKRIKOFIEG
SFSDBGTVSEEDGQSYVVPITISTVLTNMVSVIQEGFYPPKVGSFHEALLGREIMVLLSAA
IDAEYRAVLSRTRNAKNPLTTKLDKYVNNPHLQMPSESVTEREKEWVERERERIKTTDM
TAENLFRDHPYLPKAIIDGILGPKRTPTALQALQREYKRCNKFNIDIVSPETLEYFLVNNRQ
VMFSNYSVTRVLDPDSAAARFSMYVLWNALFLCSGGLTQKTNSSAVKSRLILQVFLKDMHS
LFVCQRCESGFIITKSLDTFTISLKEQSKPSMGEQELITYWKAVLDALGGGGGNNKGAENV
NGLGELMVEILSADSGLLRGGGLGGDIGFEGKMKQKREDEEVRNMHLVDKKGYYFEAAKY
VHVSKGFAALSFYLLYAAAATSNPSITNNFDRAVYLLARWGDLEKPTHNLWGNVPTDEN
TSSLLSFASFWALRNAVRARRNVIDASNTSFVGRPLPLLSAFSSKMLVDNMLKNNYVKV
ENVNREKLIWKAFREMOTESIWKTSKKAASDRNVKAKQDLIRNASIGRLIVEPVGKT
PISSIALFRSMKRSRSEDLKMGSSNNKYRLARDTKTATPRNPLSYTGKIVFSLDDLKNFSK
DSYTTMKVFPLTPLDG
(SEQ ID NO: 141)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| CAA22155 (CAA22155) HYPOTHETICAL 46.4 KD PROTEIN | 37 | 0.49 |
| O95517 (O95517) DJ1170K4.1 (NOVEL PROTEIN SIMILAR TO KIAA01 | 35 | 1.9 |
| Q43688 (Q43688) GLYCIN-RICH PROTEIN (FRAGMENT) | 34 | 4.3 |
| YH00_YEAST (P38800) HYPOTHETICAL 149.7 KD PROTEIN IN IRE1-K | 34 | 5.6 |
| P97572 (P97572) CALPAIN SMALL SUBUNIT (EC 3.4.22.17) (FRAGM | 34 | 5.6 |
| O48591 (O48591) GTL2 GENE | 33 | 7.3 |

Comments:

EST confirmation of the predicted transcript:
Nucleotides 515 to 1 of CT518: this corresponds to nucleotides 188146 to 188660
of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=1308
5'stop=1327
3'start=1383
3'stop=1404
5'primer=AGCTGCCCCGATTTTCAATGT (residues 1308 to 1327 of SEQ ID NO:140)
Tm5=58.26
3'primer=CCTGCTTTTCACTGCAGAGCTA (residues 1383 to 1404 of SEQ ID NO:140)
Tm3=58.54
probel=TGTGCTATGGAATGCATT (residues 1329 to 1346 of SEQ ID NO:140)
probelstart=1329
probelstop=1346
direction1=Forward
Tm1=68.99
score1=1.88
length=97

WO 01/38351

118/201

PCT/US00/28888

CT1003
 Nucleotide
 Genomic coordinates:
 Start: 23709
 Stop: 24300 (SEQ ID NO: 142)

Amino Acid
 MDVSSYKSTIDYHNIEDMDDLQRATYKDRMETELVLEMAKKEGRYVRSLATMDELEVPEE
 PATCYTCGYTFIRRRAPPPKRKSI FREPCAYPELLPDAPSPVRLEELVDVPEGASFFTY
 PYDDGSSTSSSQAECEDDYPPPYDPSNPQRSQVCDYCTTRQVLSSMTDHARANLIKLNK
 REKKALGLGRRNNFSY
 (SEQ ID NO: 143)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| VP40_EBV (P03234) CAPSID PROTEIN P40 (VIRION STRUCTURAL PRO | 35 | 0.46 |
| Q9Y4G2 (Q9Y4G2) KIAA0356 PROTEIN | 34 | 1.0 |
| O88508 (O88508) DNA CYTOSINE-5 METHYLTRANSFERASE 3A | 33 | 1.8 |
| Q9Y6K1 (Q9Y6K1) DNA CYTOSINE METHYLTRANSFERASE 3 ALPHA | 33 | 1.8 |
| Q23804 (Q23804) SPID PRECURSOR (FRAGMENT) | 33 | 1.8 |
| O88799 (O88799) ZONADHESIN | 32 | 3.9 |

Comments:

EST confirmation of the predicted transcript:
 An isolated EST has equence identity to nucleotides 1 to 736 of CT1003: this
 corresponds to nucleotides 23688 to 24423 of the genomic reference sequence.

WO 01/38351

120/201

PCT/US00/28888

CT1005
 Nucleotide
 Genomic coordinates:
 Start: 31091
 Stop: 31961 (SEQ ID NO: 146)

Amino Acid
 MEGEHQYLNLVREILRGVKKDDRTGTGTLSIFGPQMRFSLRDDTIPVLTTKKIFWRGVV
 EELLWFIRGNTDAKELAKKKIHIWNANGSREFLDSRGLYDRAEGDLGPVYGFQWRHFGAE
 YDTCSSDYTGKGIDQLANILKTLRENPDRRMIMTAWNPMDLHLMALPPCHMTAQFYVAN
 GELSCQLYQRSQDVGLGVPFNIASYSLLTHLMASMVGLKPGEFILTLDGAHIYNTHIEVL
 KKQLCRVPRPFPKLRILMAPEKIEDFTIDMFYLEGYQPHSGNLQMKMAV
 (SEQ ID NO: 147)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| TYSY_HUMAN (P04818) THYMIDYLATE SYNTHASE (EC 2.1.1.45) (TS) | 410 | e-114 |
| TYSY_MOUSE (P07607) THYMIDYLATE SYNTHASE (EC 2.1.1.45) (TS) | 407 | e-113 |
| TYSY_RAT (P45352) THYMIDYLATE SYNTHASE (EC 2.1.1.45) (TS) | 406 | e-112 |
| Q89940 (Q89940) THYMIDYLATE SYNTHASE | 390 | e-108 |
| P90463 (P90463) ORF 70 | 387 | e-107 |
| DRTS_TRYBB (Q27783) BIFUNCTIONAL DIHYDROFOLATE REDUCTASE-TH | 381 | e-105 |

Comments:

TaqMan Primer/Probe Sets:

5'start=424
 5'stop=446
 3'start=479
 3'stop=501
 5'primer=ACCCTGAGAGAAAATCCAGATGA (residues 424 to 446 of SEQ ID NO:146)
 Tm5=57.82
 3'primer=AAGAGCCATAAGGTGAAGATCCA (residues 479 to 501 of SEQ ID NO:146)
 Tm3=57.91
 probel=ATGACGGCATGGAATCCT (residues 460 to 477 of SEQ ID NO:146)
 probelstart=460
 probelstop=477
 direction1=Reverse
 Tm1=68.98
 score1=1.98
 length=78

CT1006
Nucleotide
Genomic coordinates:
Start: 32124
Stop: 32802 (SEQ ID NO: 148)

Amino Acid
MAFNFEDSTNLFANMDLTAGTTTDPTRPNIFFESLLPNSGIEVMKRRLVRQKCGNFEA
SGGAMSYFWLEDNAEDMENLNSGSHVKTNCLALFLQEFISNWEETDRHGQYCTFPQYMD
GGDGSRRGGYFTSLAMKWMARDVTFFVFVDRNNTVENAASIWMYQKLLAIGAKVVKVIVDN
ASNPMFVSVCNACRCKYPGPVSYVIEGHGVGHSDLTCD EISGFFV
(SEQ ID NO: 149)

Comments:
EST confirmation of the predicted transcript:
An isolated EST has equence identity to nucleotides 1 to 624 of CT1006: this
corresponds to nucleotides 32182 to 32805 of the genomic reference sequence.

TaqMan Primer/Probe Sets:
5'start=268
5'stop=289
3'start=340
3'stop=361
5'primer=TGCTTGGCATTATTCCTTCAAG (residues 268 to 289 of SEQ ID NO:148)
Tm5=58.01
3'primer=CGTCCATGTATTGGGGAAAAGT (residues 340 to 361 of SEQ ID NO:148)
Tm3=58.72
probel=CGACATGGACAGTACTGT (residues 322 to 339 of SEQ ID NO:148)
probelstart=322
probelstop=339
direction1=Forward
Tm1=69.04
score1=1.95
length=94

CT1007
Nucleotide
Genomic coordinates:
Start: 32947
Stop: 34216 (SEQ ID NO: 150)

Amino Acid
MDSNTSILPPSKRPGNLNLLQVLGIIITVALIASVSSSFIFYRVGKRKYPPSSSSSELSDV
DNGVEGGGGTTTTPTQSPDGGDGYVDLSPQKKAELRTRVANVIFQEVSKDQGVAFRRAM
NDSTDKIMEETEARINNFSEPFREATVEREVFKDDTDKNFILSTLDLTEEQFKDIVMAEV
KNQLENFDYEDMTRLIFDNIPETDYLWTHFDPKKYDTYSEKVLGFSINSIERISSTFY
KGKKYEVTTGNVAVLVDFESETIKEKAGNSLIRNVEFIVVDEQTYKSFFPAFNQVFFSFK
VNKEKREVTVSINNGCVGIVANITPLTPVGAASGHYIYGTSTAKEKTYLFVIDKYDTTE
FVCGLSNKSSTPLMALNILFMSDVFPSFDEAERPLTDAKAVEILGKRLGVGRYTNANIRN
TQ
(SEQ ID NO: 151)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q56711 (Q56711) HOOK-ASSOCIATED PROTEIN TYPE 3 | 39 | 0.045 |
| O77363 (O77363) MAL3P4.5B PROTEIN | 39 | 0.059 |
| YLM5_CAEEL (P34379) HYPOTHETICAL 49.8 KD PROTEIN D2007.5 IN | 37 | 0.23 |
| GBA2_CAEEL (P22454) GUANINE NUCLEOTIDE-BINDING PROTEIN ALPH | 36 | 0.51 |
| Q9YW03 (Q9YW03) ORF MSV089 PUTATIVE NTPASE, RABBIT FIBROMA | 36 | 0.51 |
| P87199 (P87199) KINESIN MOTOR PROTEIN | 36 | 0.51 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 907 of CT1007: this corresponds to nucleotides 33238 to 34144 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=648
5'stop=674
3'start=773
3'stop=799
5'primer=TGACACGTACTCTGAAAAGGTATTAGG (residues 648 to 674 of SEQ ID NO:150)
Tm5=58.05
3'primer=CCTTCTCTTTTATTGTTTCAGATTCAA (residues 773 to 799 of SEQ ID NO:150)
Tm3=57.41
probel=ATGTAGCTGTCTCGTTG (residues 752 to 769 of SEQ ID NO:150)
probelstart=752
probelstop=769
direction1=Forward
Tm1=68.89
score1=1.89
length=152

WO 01/38351

123/201

PCT/US00/28888

CT1008
 Nucleotide
 Genomic coordinates:
 Start: 34217
 Stop: 35048 (SEQ ID NO: 152)

Amino Acid
 MEGVILDKIETIAKRASPSYGSIDVGTAILRRQFMKIRGKINEETMEKIMGTKEERED
 TIRSIVANVIKENTVKENVTEKIRAMTDKELNDNREFMHDFGKISTGDGGTFHLFEDTPG
 FESALKAEYKNVPGATTPKYVSMNSLRIDAINGKIEEVYNPSPIMGIREYGTIRRGYEE
 NAGSKELVFMTKIEKRPNNVAENLIIRVANQQYNVMRMVFFIDYETKKGVSKEEMFIPYN
 VQKTKALKGRSTYFSFVRKIPDEPEGSIIHALGFY
 (SEQ ID NO: 153)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| BAG_STRAG (P27951) IGA FC RECEPTOR PRECURSOR (BETA ANTIGEN) | 41 | 0.007 |
| Q99051 (Q99051) IMMUNOGLOBULIN ALPHA FC RECEPTOR PRECURSOR | 41 | 0.007 |
| YIBA_ECOLI (P24172) HYPOTHETICAL 31.8 KD PROTEIN IN RHSA-MT | 38 | 0.079 |
| Q46749 (Q46749) ORF-A1 | 38 | 0.079 |
| Q25920 (Q25920) MATURE-PARASITE-INFECTED ERYTHROCYTE SURFAC | 35 | 0.53 |
| Q06166 (Q06166) MATURE PARASITE-INFECTED ERYTHROCYTE SURFAC | 35 | 0.53 |

Comments:

TaqMan Primer/Probe Sets:
 5'start=361
 5'stop=384
 3'start=429
 3'stop=451
 5'primer=TTTGAAAGTGCTTTAAAGGCAGAA (residues 361 to 384 of SEQ ID NO:152)
 Tm5=58.35
 3'primer=TCGCATCGATACGTAAACTGTTC (residues 429 to 451 of SEQ ID NO:152)
 Tm3=57.96
 probel=CCAGGAGCAACTACTCCA (residues 397 to 414 of SEQ ID NO:152)
 probelstart=397
 probelstop=414
 direction1=Reverse
 Tm1=69.05
 score1=1.94
 length=91

CT1080
 Nucleotide
 Genomic coordinates:
 Start: 291719
 Stop: 292205 (SEQ ID NO: 154)

Amino Acid
 MTSPAPSPSSTPKSSCTTIVNRCGFLDNNKEVVIYDTNSKFKCEPKNLELIGVLSGVSD
 NVVTQISPDQIFVGTVMVKYNWSKSGHERFSDMSNNCLDNITRPSEVIESVIKKTSSDFK
 MKYTRSLMDHTEKYFSGDQKLSKISSWCTTPIRQWVCNSV
 (SEQ ID NO: 155)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| CAB45519 (CAB45519) GIBBERELLIN 20-OXIDASE-ARABIDOPSIS THAL | 32 | 3.0 |
| Q52743 (Q52743) CELB | 32 | 3.0 |
| Q39110 (Q39110) GIBBERELLIN 20-OXIDASE | 32 | 3.0 |
| O60963 (O60963) L549.2 | 31 | 3.9 |
| FLPA_ARCFU (O28192) FIBRILLARIN-LIKE PRE-RRNA PROCESSING PR | 31 | 5.1 |
| DPOE_YEAST (P21951) DNA POLYMERASE EPSILON, CATALYTIC SUBUN | 31 | 6.7 |

Comments:

TaqMan Primer/Probe Sets:

5'start=201
 5'stop=220
 3'start=290
 3'stop=312
 5'primer=CCCCGACCAGATATTTGTGG (residues 201 to 220 of SEQ ID NO:154)
 Tm5=58.70
 3'primer=AGGGCGTGTAATATTGTCCAGAC (residues 290 to 312 of SEQ ID NO:154)
 Tm3=58.09
 probel=CGCTTCAGTGACATGAGT (residues 265 to 282 of SEQ ID NO:154)
 probelstart=265
 probelstop=282
 direction1=Reverse
 Tm1=68.90
 score1=1.90
 length=112

WO 01/38351

125/201

PCT/US00/28888

CT1009
 Nucleotide
 Genomic coordinates:
 Start: 35073
 Stop: 35967 (SEQ ID NO: 156)

Amino Acid
 MALQEKDITIGNVSAALRELMYSPTHMQHHDKLNTFLDRNVESSEEKIRQIVDKIRSQT
 TSDISETVNNVTNGTAFSLFEDTLEGVMVKKNIGDNLQSGDFIDGRKKLNDMKSLATGAI
 LSRQRDFVAESITGTKDWLKAIMGCGIIRYTVFVNNLARSTLDNDDKAATYYNTPIYGG
 YCKMAIKDYEIPDSYSKVEAHTVEGRKMTFNIKWRGDTINNLITIIPSVTGYLASISED
 ADVQAPLLLNCNNCFIEADMSSLYMDEKKTEASFTLNLPEIEGADANAVEICIVVV
 (SEQ ID NO: 157)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| YB75_YEAST (P38321) HYPOTHETICAL 101.2 KD PROTEIN IN FAT2-P | 34 | 1.3 |
| Q08281 (Q08281) CHROMOSOME XV READING FRAME ORF YOL138C | 33 | 2.9 |
| Q92271 (Q92271) 12.8 KBP FRAGMENT OF THE LEFT ARM OF CHROMO | 33 | 2.9 |
| Q26032 (Q26032) VARIANT-SPECIFIC SURFACE PROTEIN | 32 | 3.8 |
| O02179 (O02179) CALYX PROTEIN | 31 | 8.6 |
| O01394 (O01394) POLYHEDRAL CALYX PROTEIN | 31 | 8.6 |

Comments:

EST confirmation of the predicted transcript:
 An isolated EST has equence identity to nucleotides 1 to 780 of CT1009: this
 corresponds to nucleotides 35166 to 35945 of the genomic reference sequence.

CT1081
Nucleotide
Genomic coordinates:
Start: 292189
Stop: 298777 (SEQ ID NO: 158)

Amino Acid

MQLRLNENFVKEEHETVVVHNPSGMTGFNIFNSSPVYFEVHNEMDALIFMAAFLKHNSLWG
EINANMDLYTFDYAGAFLLDERWCHHEKSFSVVRQLINSYKCRRKIMQALDNNYNNKNK
KRKNVGGAPAFTFMSGDGEKGKEALEASFDVIGGTRGGREFGVDSTPCPHSSAMQLKLDNE
GNYGCIACFASMFVLENPGDESSFISTDASKIGQAQAWIDERLRNNENGGEENNFKKT
FHMLADITQKAHETAYSNTIPLGPNGRQWNWPTHVPEIAHEFVTHSLVNTLKNLGDRL
PRENFIDILYNLLNPFQKMLLVFIQNCILTGHKNNENNVPRGSASGKWWTINFGVNMWT
FQVTKCKVEKDRKISDLACMETLPLPNDPGSTTVDDRIVFKGFCRGENLGSVGEVVDIT
QSVKNFCLMVENRKFESVDKETGFISSESIVSDPFFSLEVTGCRSNRAQDTINNGRV SARV
MRILKSREGARVWLAKDENAIIFENVNHDIAISTDAMERAIGQHKILYDIETTDKDFTD
KKSIVTSIGFCLCTGGDMTHGGERGVFGLVAPGSDVEKVKETIINSYDPEEKEDIMKQCP
QVIEIFTNEFEMLLGFGKYIDKVKPHVISGWNNVAFDDPFVTRIVKHLSDHTKDMSYCV
ADASTAESVLPRAEGGGGGGETPYRLSTPQERIQLASTGIFNKLKGFVDKKTGMLKPEMT
ADLLAGAESQANTKFERNKLSSSNKGSAGWFQKIIGGMCSAIRLDLMKVCEKAYKESLS
EENLNAVLAKVSSVGDVKVNVKDEVLDHFHLLGLFLKLKKAQDQAKVHVYCKDAYLTGIV
STSINKEGEIFRLCMSALTEAVVTANLATPLCIGEGAICRNMGEERADRRGVGVRRHSI
ATDTKGMVSQPIVNHVPYQTDIMTSLYPMTMCQNNLCTTTFVTHRQIMQLRDLRVLEKM
KNKTTDSLILLDVIDECNQIVLSEYRPIDIAVASWKNSNSNRQTPITRIEESLGLRFIEN
LDAEKTNNKTWCTNTSPNMNVTAAGMDYFPEIVCDINMQFAAKVNDDMHIAFASLEYMLQ
VLPIMLIDRPYIGAHITAGKCRTELDILSELEKDFSVEKDEEIIRTHWTFKGQKQYDFCH
SPVTQMARHIIESTGRNIRDYEGNEKFERLVSLSDRIYRRVGAFDSANDPAVRLWSSRLI
NVGMLVVRTWNVKTDILKGIIPQMQATYRADRVVMQNKAKEFAKMGDMKRAGLNKVGQNM
KLGMSMYGHLALRARSSRKEFASGSANTASSISNMSATGGIGGGGTRHSVTANQITENAR
CVFEGNIGCGLOMALPGTKQTYGDTDSVFCVHNIVGDGGMIPEYDEQTKGYYVMDIALKN
KMAAIIPIVLVNSLTGIGIQFVERRDAGVGMNIAHERLAVAGLLFAKKTYYHMLHFNENSAA
FNDMIKCLKSTDNNNKFASFIRKPSHADGYVVPHPNPSLILRAAEGPAGKKLKSFLLEEGLH
DEKSMEEWFTSSPTWMAMDASVINNYASQIVGVEKGNWIDAMTSRPIEAGTEMMEAVTQ
ANAAFTPYKKGAFVKKGITPTTKLKLGLQSLIARFLPKIEEKKSCYLDVMKNHVENFASHI
TNPAMMITSSRVNKFDTSKESQSRPNPLALAINNHLNPSSEISLGQKFKTVTSVSSWSLSA
EEGEVPAGYFNAGSVRW DATNMKGSVPAPFSVKNLSVVPNAITSVYKMMVESDKTAIKSMIA
KNVEVLCSTANTGFSLRGALSFN TGVI VTKDVAMACIRSLNNQMMLLFVGGGKDYGED
DDDDDEEAEEDEENGENEENKGD CVTEKKIPGRSTNKDVGEETKTSEKTEGERKGSKTA
KGKTEEIIASSLSKCGKKDARDVILDRLLKATHSSCTNNEERTRVLQYYSNCTLSYYITSV
MKLDQRVADQMENLISQLDQIRNLSNKKRQEKGGPFKSELDAMVAVKVFFPVLDASRK
LTQDHWWKCPVSI PETREEKPLMGVPFEVALNSLIGKHKCTDTCDMACQSLYFVLLYTL
ALKFENERLARQIGLDDSVDLMAEMLFGGDKLLAQEVLKRVKDAQDRKLVKSLPLNLYNH
DTNTIIFLFESLRFAPKPVAGMSVSEIKDAVRGLAFSTTGTGVWNYTDERFFGPLYNMDE
LCNERVNGNCKLSFITGIYHTAAVELAAACLSCVL
(SEQ ID NO: 159)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| DPOD_YEAST (P15436) DNA POLYMERASE DELTA LARGE CHAIN (EC 2. | 52 | 5e-05 |
| DPOD_SOYBN (O48901) DNA POLYMERASE DELTA CATALYTIC CHAIN (E | 51 | 8e-05 |
| DPOD_SCHPO (P30316) DNA POLYMERASE DELTA LARGE CHAIN (EC 2. | 50 | 2e-04 |
| CAB58156 (CAB58156) DNA POLYMERASE DELTA LARGE CHAIN | 50 | 2e-04 |
| UBF1_RAT (P25977) NUCLEOLAR TRANSCRIPTION FACTOR 1 (UPSTREA | 50 | 2e-04 |
| UBF1_HUMAN (P17480) NUCLEOLAR TRANSCRIPTION FACTOR 1 (UPSTR | 49 | 4e-04 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 4 to 1205 of CT1081: this corresponds to nucleotides 297608 to 298809 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=2738

5'stop=2760

3'start=2817

3'stop=2838

5'primer=TCGTCAATCATGTTCCCTATCAA (residues 2738 to 2760 of SEQ ID NO:158)

Tm5=58.05

3'primer=TCGATGGGTCACAAAGGTAGTG (residues 2817 to 2838 of SEQ ID NO:158)

Tm3=58.56

probel=GATGACCATGTGTCAGAA (residues 2787 to 2804 of SEQ ID NO:158)

probelstart=2787

probelstop=2804

direction1=Reverse

Tm1=69.00

score1=1.99

length=101

CT1082
 Nucleotide
 Genomic coordinates:
 Start: 300935
 Stop: 305108 (SEQ ID NO: 160)

Amino Acid

MTEQGDQGIKVRKHLHGPRGERGETGPAGAVGPAGPQGERGAIGPAGKDGAVGPAGPQGER
 GAIGPAGKDGAVGPQGPGERGENGRPGRDGAVGPQGERGAIGPAGKDGAVGPQGERGAIG
 GPAGKDGAVGPAGPQGERGENGRPGRDGAVGPAGPPGERGAIGPAGRDGAVGPAGPPGER
 GATGIPGRDGVDSVGPQGERGEIGRPGRDGAVGPAGPQRRGATGRAGKDGAVGPAGPQ
 GEKGEAGKDGSIGPQGIQGPGETGPPGRDGTAAERGERGFPPGETGPPGKDGVDSGSE
 GPQKRGETGPVGPGEPLAGLPGRDGAIGPAGPPGERGATGLPGRNGVDGSIGPQRRR
 GATGRAGKDGAVGPAGPPGERGATGIPGRDGVDSVGPGERGETGPAGRDGAVGPAGPH
 GERGENGRPGRDGATGPIGPAGPQGEKGENGRPGRDGATGPIGPRGETGAMGKNGVDGSM
 GPQRRGATGRAGKDGAVGPAGPPGERGETGPAGRDGAVGPAGPQGETGLTGSPPGRDGAT
 GPIGPAGPQGEKGENGRPGRDGATGPIGPAGPQGEKGENGRPGRDGATGPIGPAGPQGET
 GLTGRPGRDGATGPIGPRGETGAMGKNGVDGSTGPQRRGATGRAGKDGAVGPAGPPGER
 GENGRPGRDGATGPIGPAGPQGETGLAGLPGRDGAIGPQGEKGENGRPGKDGTGPMGPP
 GERGETGPIGPAGPQATGLPGRDGVDSVGPQKRLIGRTGRDGAIGPVGPAKGET
 GLAGLPIDGKDGSVGPQGAIGPIGPRGERGETGRPGRDGEDGSTGPMGPQGLRGATGAP
 GPQGERGLKGRPGKDGETGPPGRQGRDGIMGPRGLRGEKGA PNGDLEGPEGRDGAPGPA
 GPIGPQGIKGLKGIQGRPGRDGEMGPAGKDGIEGPRGQDGTGAKGPRGLRGFGRTGET
 GAQGSRGEKGDRLTGPPQGRDGPPEEGPQGLRGERGAPGPRGPRGIRGRSGPQGSNGVQ
 GPRGPRGTGRTGIQGLTGIEGPRGPRGIQKKEGRMGKIGHRGEKGDGDRGEQGIAGAD
 GEKGPRLRGIRGPIGAPGKPGTEGVGRGPRGVRGVPGYPGAQGELGPQGPPTGPQGPAGPQ
 GPMGRTGDTGPMGPPGAVGPRGEKGGRRGKNGKPGADGKDAVNIIQKYSITHARAEIM
 WEGNEIGEAYIGRSYGTDTIPVMIENRIGMTNEDKKNEYCIQVMTMHSITTRGRTSGVFFV
 VSNKTDYILLVTLMPESVSCRTDVS TNARSERVNAVRESKSYRFRPSDQSIGTHSR
 SKIAVVMYPDASMSYSVDTLADVARRETTSVLLLAETIHGEKDRGFYADRGTVGRLMVP
 PTEEEELLVLQX
 (SEQ ID NO: 161)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E- Value |
|---|-----------------|-------------|
| Q14054 (Q14054) COLLAGEN TYPE VII PRECURSOR | 940 | 0.0 |
| CA11_CHICK (P02457) COLLAGEN ALPHA 1(I) CHAIN PRECURSOR | 935 | 0.0 |
| CA17_HUMAN (Q02388) COLLAGEN ALPHA 1(VII) CHAIN PRECURSOR (| 933 | 0.0 |
| Q63870 (Q63870) TYPE VII COLLAGEN | 928 | 0.0 |
| Q60444 (Q60444) TYPE VII COLLAGEN (FRAGMENT) | 925 | 0.0 |
| CA11_HUMAN (P02452) COLLAGEN ALPHA 1(I) CHAIN PRECURSOR | 923 | 0.0 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 557 of CT1082: this
 corresponds to nucleotides 304552 to 305108 of the genomic reference sequence.

CT1083
 Nucleotide
 Genomic coordinates:
 Start: 50094
 Stop: 50295 (SEQ ID NO: 162)

Amino Acid
 MAVTEIPCGTRNIAEEDVELELILVTAEAEVREMAAALAAAIIIGAVVVQIGRVLDEVVA
 AEVELM
 (SEQ ID NO: 163)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| LIPB_MYCTU (Q10404) PROBABLE LIPOATE-PROTEIN LIGASE B (EC 6 | 30 | 1.8 |
| Q64033 (Q64033) ANTIGEN LEC-A | 29 | 3.1 |
| OL56_STRAT (Q07017) OLEANDOMYCIN POLYKETIDE SYNTHASE, MODUL | 29 | 5.2 |
| RPSD_PSEFL (P52326) RNA POLYMERASE SIGMA FACTOR RPOD (SIGMA | 28 | 6.9 |
| O54540 (O54540) RNA POLYMERASE SIGMA FACTOR | 28 | 6.9 |
| Q9YAS2 (Q9YAS2) 136AA LONG HYPOTHETICAL PROTEIN | 28 | 9.0 |

Comments:

TaqMan Primer/Probe Sets:
 5'start=24
 5'stop=45
 3'start=95
 3'stop=112
 5'primer=TGGTACTCGGAACATTGCAGAA (residues 24 to 45 of SEQ ID NO:162)
 Tm5=59.21
 3'primer=CTGCTGCCATCGCCTCTC (residues 95 to 112 of SEQ ID NO:162)
 Tm3=59.19
 probel=TTGTAACAGCAGAAGCAG (residues 71 to 88 of SEQ ID NO:162)
 probelstart=71
 probelstop=88
 direction1=Reverse
 Tm1=68.98
 score1=1.92
 length=89

WO 01/38351

130/201

PCT/US00/28888

CT600
 Nucleotide
 Genomic coordinates:
 Start: 1118
 Stop: 491 (SEQ ID NO: 164)

Amino Acid
 MHMWGVYAAIAGLTLILVVVISIVVTNIELNKKLDKKDKDAYPVESEIINLTINGVARGN
 HFNFVNGTLQTRNYGKVYVAGQGTSDSELVKKKGDIILTSLLGDGDHTLVNKAESKELE
 LYARVYNNTKRDITVDSVSLSPGLNATGREFSANKFVLYFKPTVLKKNRINTLVFGATFD
 EDIDDTNRHYLLSMRFSPGNDLFKVGEK
 (SEQ ID NO: 165)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| 051578 (O51578) EXODEOXYRIBONUCLEASE V, BETA CHAIN (RECB) | 39 | 0.033 |
| 068195 (O68195) DIOL DEHYDRATASE-REACTIVATING FACTOR LARGE | 32 | 3.2 |
| BACC_BACLI (O68008) BACITRACIN SYNTHETASE 3 (BA3) [INCLUDES | 31 | 5.5 |
| TKT_MYCPN (P75611) TRANSKETOLASE (EC 2.2.1.1) (TK) | 31 | 5.5 |
| Q9ZER8 (Q9ZER8) REPAC1 PROTEIN | 31 | 7.2 |
| CAA10001 (CAA10001) REPAC PROTEIN (FRAGMENT) | 31 | 7.2 |

Comments:

EST confirmation of the predicted transcript:
 An isolated EST has equence identity to nucleotides 1 to 569 of CT600: this
 corresponds to nucleotides 513 to 1081 of the genomic reference sequence.

CT520

Nucleotide

Genomic coordinates:

Start: 209342

Stop: 196799 (SEQ ID NO: 166)

Amino Acid

MFKANVLNLGGGKFLESVDVRDHLIKCANQMKEEPTTLRICLSNKLPEYDNRRLPLLLLNE
 GEQILVTDNLTKNGNPLVKQMGHLAVQDRVGGDGSVNPNLLYAGCNVVEYDTVNRGNDG
 KLIMYSQPATLKDVAKSKKKGMVKVKVPEITGDQFLDKLNERSCQENRRMDEEGPHVG
 TGKLLRELIIMRLYEETSSAEKLCVTPAFREFLGCGRTATDVPVFKVAFITNASLMGL
 KVIFYPTILEEERLAAVSDTENVVLLKSILKVOLELLSECMPRIVERVESMIKKTVACFK
 IDIGGSDNWNLPGHCKVSDTAFYPYHHAQLVGEKKNILSISNENMVTSLGVVKADRAEWM
 CKTLESFEKKCLYLENLMGSMANTDDWRRKILFSELGPEMPYRNKSLIMDQDFCTIGMCY
 KFLAEGGGLLTKTNATLLKEKMACKGLDDSGDGDDEEDNEEGSGGKSGGSGDENNI
 NKPPAPKQIPPLAANVYNSIINDDKLDQIVCFKRRKHGFFLSIDIDNSPLLAMEFLLPQ
 KAMSKKNCVERVKPETKNIIRNLTVGNTIKFDTIMPFALQIVVRYENRNLKLPRTDIL
 QQRLKNNTWDALSKGKFAEMWQFTHKESLKPTTIEELESIPPPPTQSEEEEEAAAAAAS
 STTPDMVSSLEEGATSTSSSDENQIASLENIKKLLSIITSTFATGADKNDTIFAWTVVTL
 AERFCALYNITSHPEEYYQIIREDFEFEGGFEEKFRHMCDAINRELSIYVPKSVLEKQSV
 CRMGVAAYENSMEIRIKNKTNSKLCKIKYDESTMVYELNNDTFKTFDYDESDFSFGPMYEC
 APMETFQRLFASVKSDEAVLADKKSEKREKLYQQKQEYLRKCDNDVVSARQILNNVASN
 ESDEESDEESDDEENYGAAGGATGDYGGDDEDDCYGFLGEFGSSDDENVPDNDASSIN
 NVQDDVFRDVFNTKTFNFRSSSLCHRQKYVSTVIVEEMEKNLDCVLTLDNSAASGDIKE
 INRRSLMRNWNVVPFTMPVREIVKPNVNSDGTANSNNNIPFFCSCASLNNFKSDSPLSS
 NNTMSNEKICKLLPISSSKHLKDLTVLRNTMACERRYFSDVTAALGFVKDKVNGNIR
 SILDNKRWDIAKQCKLAGKCLSSALPLGIYENVISEDNKLINTFRPRSLARLACSSGGDG
 VSDKSVNNGFFSGIWLALCANQDLESVVLGSTVVDPLKPTKVFNQSLSEKELKEKQOMCL
 DAANYFKDHNVSCLNIYECFKMMECIRMALNGKTSNDSEFFSNLITRYGSGTNSPASR
 LWTILETVRECFNNSLPIDWGSVLKDWGSDMLNLKAGVSNVDESAGVFESEFLGV SAR
 AFFGKDLDTNLDADTWECLLNDNDKDWKAQVAKAYEFALKDNDIRSVENFINSSNLLTNN
 NVIKKLLKIKPTSPNDVRHQIWEDEYYPRNKSTLRSRAEWMAATEEVLKTEMSLSCVLAM
 VAMYRIMMQGESVREIATAPRLSVDKMVPLIRCFKITSKWCSTGKGDSPKKADASIKE
 GRFYDIEEDPLHFYRFAAYVIGQVASNDIVIEEMTRKILMSFDFNGFDTSNWLQFITYRF
 SHVLGRRSRLLSRPLSLVKNLVSVSSLADKNSEKSNMYEKRVGKVMKRIARLVLVKAA
 DSVRASSNDLLDCCILDVNDVSVKSLDEFRAKTRQELQETRIDTNYNLVNSCTTAQLAA
 VEKSSRIINTNISFHNIPAGQAKVMDANEEAFIDPSLEEINKEDNSGAKQMTGKGGSNRG
 RSKKSGGGGFNNAGGFYNDSSRGSSSVVDEDSRSRTGFSQIHMDARNEEDRESGLFSYD
 GYVLNRIKMNITQONINNDIVKVISDIENFFKICVPFSKKEYALYGVTTETALSAGMDAIE
 RWNKAVEEETNKIRKECRDLTDGVSVDMMNIIICPGDYMSVGEVGGNGGCGGSSSSGHL
 SNNNNEANQTNEISEDQLKHEGSDCSFWFNFKVNSSEKKQKGKSVLANTGHEGRIVG
 RPLRTFIQYKKGFAETKVLTRYFSNHISDSYWSQVMPICYIKNMALGDEDKSKKKFGKR
 PWKNFNNSNSSSNSVSVKYSIQDLEKKDSLKNVPMGYDEDLLSLYDDSLTSTEKLNI
 KIVNDSKDAYVILGSSNQSSFDQTFSSQYFTHQKISNINTYKSLGKMWNCCNMGSPKNQI
 VLLKLLLFKNLNLWIKLYERHISVLCNWGCIHPNSSKNSHFEMTKNNAPCGVTDSPNPL
 SVYHSGFLSVEDYQQLKDTFPLMNLHRTFSAKSKDNNSSDPSPEKISAASLAKAVYARE
 VLSSCLDPEGNFCTSWITNSCSVLFTPGTNIRRGDFFNKSICYRQDNDYCFIGKEETKK
 CPNFVSSEIEIVSILKTAVFLSTNSDGHKRVLRVINYNKDHSGLYAGIDTGCADDEDDDD
 DQGGTDKTCLLQEDSMDAKRMLISMRSVINGKSLDESSLAIKDNFNFLAGTDKGFYLDN
 SFFNSPVQGFVAPRGTKIFKKCCDFLLNKGTGGVFARIFFTDWACIVSSSKGKNNKAI
 ESTLQIRNGGCFSERLTPSMFDNESEQGELFHDRYCPDFLSDYNKQNIIFSEQAYKCSFLA
 NPVCPAKNMLKRAKNIRLCITNAGTALISKIMAEVEKMGNARTFISNGTAIPFRLAENTA
 CISVDNNRYFLIDGTYLLGGRLEGINLVTDMYTRCKLKAKEHVILNSLSTEFISAALAS
 SMEGTTMGRGLCLIEHVSVMKNTDSVSNMKNFWSMAEDQEETDENEDDDDENEEDEDEN
 EENTENTSUVKYEVPVSKTAFSSSLKPPSIFIADEYIFLSILYELAKATSDCETASSSSS
 SSSSSSSSKHSSSSSSSNKKRKQKDDVNSTTTALHALRKCYISCVQKTGMPRMDVVYL
 LRGLMNFEGGMCTAIASGDGEKAHMQVTLCSVALNIATKAVVFGTKGNNLKTTLVDLC
 KRTWFERFTNINVTALNNAGDSSSTQANLASFAGKKGIVIIDEVGHQGSFGSKKSSSED
 DKDESASRSNGVDGSGSGGEMNSVDINEARNAYDGGNSKIVFSNINRLMTESKLKVCDO
 EYDFISSELKHEKNRNACNDTKKRKRGEIEDEGVECEEIERNDGKNDENGVRIKDPINI
 SFFARKAHWWNCSSGVVSTTFKEKNIVYNMLHRGAMPFSIKDCTDSPWLNETDAVYRHCK
 KPIEYEGKFSKSEVKTALKCILGKFGSKICDNESFESIIDENCQVNNLHWNDCEDIDE

WNEKFMSKNKKNKQNMKIEDKVDAIMNIIQKNNGLLKWNTSFDRDGPVLVCNPATERFS
 EMITSSLSAQDMLEIKKYLGDNCLSTNGGVKKSVIDGNTSAPGVLIAYHCVYTGKISDDL
 SKTNPVLLPPPKQHFAVDDAAEKALLGPTLSNINIDSIRNIKTISRKLSSIIKDPEA
 AKLLVDRDLDFMNMERYDASLFDVVKPSKYSFPGFTSDGSVVLSTSTSDCENVLSCLK
 KRIEKDKMSAKNSGSFIRMCMDKNLLSDEKDDSSSNSSKNTSSLPKTDNSSDIANFLSV
 FGENRQSSQFSFASNSSGGGDSNKEACFNVDTPKRRQLVSALQKHNSDGSSSIITEIAK
 AIPQKNDVSSSITKHMLPGQFPSSLLKNMTSPQNSVMIRGIFQQGAKSSITVSPIMMSNS
 YIFSFFVDEAMSKRLIVFPCDTTFVFENKNEDVKKIIGLLDRGMKYIHSSLMMERCIKFG
 KHGIKQRQHEFNHKKAWNDFSGHSSDNKKKDRISDVSSVLPVLMKNLIRNKVLELRDV
 KSVSRLEENTNTFFHLYTSMMLCAKAATNYGESSSSSATITEVEEDNSCDAEEQLRRKK
 PANYESMCNKLPSPLQMCQINPKSLNTMAMNIARSQGAQAQLNSMLNSVLFVEMPFVKT
 TRFFGRDFNIMHSPATKNRPAINFDNCIGMSLPNPDMDVVGYDKEGELIGVGSSLTKHL
 CDAGWSMDVDRDLMSCHHLHMLFEMALQYTECKRRLSSLKTLKSDKTGVDYVAVMLACMV
 YQLMVSNLKYPVFLSSSSHKRANTEDIADENQVSSLSVPMFLAMVVNKPLHALRHSTNLA
 LPNASQKSDHSDIVKYIVMNQWGLRLNPDYLCPCNVKHVL
 (SEQ ID NO: 167)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| SR40_YEAST (P32583) SUPPRESSOR PROTEIN SRP40 | 54 | 2e-05 |
| Q9Z1T1 (Q9Z1T1) AP-3 COMPLEX BETA3A SUBUNIT | 54 | 2e-05 |
| AAD56625 (AAD56625) NUCLEOLIN-RELATED PROTEIN NRP | 52 | 7e-05 |
| P90493 (P90493) HERPES SIMPLEX VIRUS TYPE 2 (STRAIN HG52), | 52 | 9e-05 |
| P70475 (P70475) NEURAL ZINC FINGER TRANSCRIPTION FACTOR 1 (| 50 | 3e-04 |
| O95815 (O95815) DENTIN PHOSPHORYN (FRAGMENT) | 50 | 3e-04 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1064 to 1 of CT520: this corresponds to nucleotides 196778 to 197841 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=7375
 5'stop=7397
 3'start=7450
 3'stop=7472
 5'primer=GATGACGATCAAGGTGGTACAGA (residues 7375 to 7397 of SEQ ID NO:166)
 Tm5=57.66
 3'primer=CCATTAATGACAGACCGCATAGA (residues 7450 to 7472 of SEQ ID NO:166)
 Tm3=57.78
 probel=CAATGGATGCTAAGAGGA (residues 7424 to 7441 of SEQ ID NO:166)
 probelstart=7424
 probelstop=7441
 direction1=Reverse
 Tm1=68.98
 score1=1.98
 length=98

WO 01/38351

133/201

PCT/US00/28888

CT601

Nucleotide

Genomic coordinates:

Start: 1511

Stop: 1196 (SEQ ID NO: 168)

Amino Acid

MYLSHIRQTPLVEERRALTFKMYHHNNNNQHSFVNCQCRRITSSSINCSSCSRETFNSVKA

IQYFNKTSRNNTAHHFKMPASKDRNYSSFYAETAVAAHNISQW

(SEQ ID NO: 169)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| YT66_CAEEL (Q11082) PROBABLE G PROTEIN-COUPLED RECEPTOR B05 | 32 | 1.4 |
| O00885 (O00885) MAP KINASE KINASE PROTEIN DDMEK1 | 30 | 4.2 |
| CHD1_HUMAN (O14646) CHROMODOMAIN-HELICASE-DNA-BINDING PROTE | 30 | 4.2 |
| O97292 (O97292) PFC0965W PROTEIN | 30 | 4.2 |
| O96226 (O96226) SER/THR PROTEIN KINASE | 30 | 5.5 |
| O96563 (O96563) C-13 ANTIGEN (FRAGMENT) | 30 | 5.5 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 353 to 3 of CT601: this corresponds to nucleotides 1184 to 1534 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=96

5'stop=116

3'start=142

3'stop=162

5'primer=CTTTGTGAATTGCCAGTGCAG (residues 96 to 116 of SEQ ID NO:168)

Tm5=58.03

3'primer=AGTTTCACGGGAACAGCTTGA (residues 142 to 162 of SEQ ID NO:168)

Tm3=58.43

probel=TCTTCCTCCATCAACTGT (residues 124 to 141 of SEQ ID NO:168)

probelstart=124

probelstop=141

direction1=Reverse

Tm1=68.98

score1=1.98

length=67

WO 01/38351

134/201

PCT/US00/28888

CT521
Nucleotide
Genomic coordinates:
Start: 272387
Stop: 268691 (SEQ ID NO: 170)

Amino Acid
MSKSSSVTKSASFNSLMENAPSSKIELLEDGWTKKAAAADTDPTAKPTGLSISLMDI
SGSMGSKSAVADSCSGIMATLNVIAPGIQNAIVYYNDFDKHSIESGPVVRAPDCSEWEG
GDFVKHMRKTEVCGGGGGSEALHSSLMYVFNNMIPAFKKMHGITRDEKFPILIFVFTDE
DVRIANSDTGKLCANSYDSETAPEEEFIMKTWGQKPLTILDMRKALVENDCWLRLNFSR
CSGSNQSELQCEDVINFSGYDNNRWQLFESFDRRSCNVRKNIATFIMRQISLFPKNLNDQ
FSAFPILREINQEELNVFIESEGRSEPAFGEKYGDAQRESFKSRVLNMAPLDFGRVVQGG
GRYNNHKRSVFLNCAYDSAFCCSKQTFNPPQQQQQQSSSGGGGISKLA VVTORASITG
GGNAASTLALHMNACFQSLDDFGIDHTNLCDCKGCTKLMAVEATSDQGRKTKLSRKYAR
VHWAKMFAEKLFKMMIKEQSMYACSAVPDEIGAIYAFVTGNNAGVCSRSTILSDLGTE
CGNKA EYAFLEKEGKHKMSASYDALQVINNTDLTPEQSSMFMWFYVPNDAL EEA GKI F HQS
FSFSNSYTGGLLSLDEYKRFEFGQCFDFIKKLVSLKITRNVEDVLL ETSKTSNRYFAI
PVFCGSDDQKEVLREELASDLFGGREDVAEMMFIDLETVIQKLGTL YDVRLSLPEGGYAA
IKSVCAAAASWAASCEVPSNTSNMILSI AKMAFTKYQE QNSSSETDLDIILPSILEGTAD
GEIENNL SGVFLRCLITWANKIGVDKNFTNKLEHFLALRILTKAGDSKIGEKYETFPVR
RLDLSEKDLKYICKRCGVKSLKMEYDNDEKLC LCKGN YRMGKPMVYHWDNKLTRDPRAK
TASPTTLNLNAKKIDDKVKEMASDIIGALNLPPTDKDNEIAVSAAAKAVGILYGKTCLL
YKLLNEGNIDIPVAVCVECDCCKSKYMMSTLGPDKPQNRKCPWCYANKLVAMGRGGKKL
LMDLIECGAPSLAMVEEAIRTS GDV MYEELGEGEEFYIIDYFLKLKNTAIAEGNKLQNN
NKRPAFLQVTSPPSPPKMRS DLPDSLLAAIGECAIETKEKTTVNLI GLGEVKVVENVGP
NDLDGKDPFISLQEYCSWDKFNSLFVNPWLG YRLDEQWDDWNTFLIHVKKNDVWKFLCNK
TSPFSVVVMNDGSGLLNVDNVNVLVRQKICV
(SEQ ID NO: 171)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|--|-----------------|------------|
| AAF04634 (AAF04634) HYPOTHETICAL 45.2 KD PROTEIN | 76 | 1e-12 |
| Q9Y1T1 (Q9Y1T1) DNA POLYMERASE ALPHA CATALYTIC SUBUNIT (EC | 36 | 1.3 |
| P91805 (P91805) ARYLPHORIN GENE-SPECIFIC BINDING PROTEIN-2 | 34 | 6.4 |
| O45322 (O45322) DY3.5 PROTEIN | 34 | 6.4 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 895 of CT521: this corresponds to nucleotides 268736 to 269630 of the genomic reference sequence.

CT602
 Nucleotide
 Genomic coordinates:
 Start: 2996
 Stop: 2702 (SEQ ID NO: 172)

Amino Acid
 MDILEDIYKSAITLVLSPEFVNDVKQEASQVVEGLIPSIREAVFRRLLLEEERKKHEDEV
 GDVEDKRQAVIDKANTMITTMAAEYLESVDILEEFGF
 (SEQ ID NO: 173)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| CRCM_HUMAN (P23508) COLORECTAL MUTANT CANCER PROTEIN (MCC P | 36 | 0.055 |
| ACVS_PENCH (P19787) DELTA-(L-ALPHA-AMINOADIPYL)-L-CYSTEINYL | 33 | 0.48 |
| ACVT_PENCH (P26046) DELTA-(L-ALPHA-AMINOADIPYL)-L-CYSTEINYL | 33 | 0.48 |
| Q9XJCS (Q9XJCS) HYPOTHETICAL 26.4 KD PROTEIN | 32 | 0.62 |
| NOSZ_ACHCY (P94127) NITROUS-OXIDE REDUCTASE PRECURSOR (EC 1 | 32 | 0.82 |
| EZRI_BOVIN (P31976) EZRIN (P81) (CYTOVILLIN) (VILLIN-2) | 32 | 1.1 |

Comments:

TaqMan Primer/Probe Sets:
 5'start=121
 5'stop=144
 3'start=200
 3'stop=222
 5'primer=AGAGAAGCTGTCTTTAGACGGCTT (residues 121 to 144 of SEQ ID NO:172)
 Tm5=58.14
 3'primer=TGCCTTGTCTATCACTGCTTGTC (residues 200 to 222 of SEQ ID NO:172)
 Tm3=58.14
 probel=AACACGAAGACGAGGTGG (residues 164 to 181 of SEQ ID NO:172)
 probelstart=164
 probelstop=181
 direction1=Forward
 Tm1=69.14
 score1=1.85
 length=102

CT603

Nucleotide

Genomic coordinates:

Start: 24906

Stop: 24660 (SEQ ID NO: 176)

Amino Acid

MTCPEISKHISGTDRRFWNTADPGGLSYFNPFLFTLHLHLKNFSKIFSAHSSLGGGPLTR

PYVKFEGWTAGSTQRQITERS

(SEQ ID NO: 177)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|--|-----------------|------------|
| CAB49803 (CAB49803) HYPOTHETICAL 69.0 KD PROTEIN | 31 | 1.8 |
| O74312 (O74312) PUTATIVE TRANSMEMBRANE | 29 | 5.2 |

Comments:

Hit to public SBV sequence:

gi|6165655|gb|AF099142.1: CT nucleotides 118 to 231 match nucleotides 5652 to 5539 of the public sequence with a 90% homology, a score of 139 and an Evalue of 5e-36

CT523
 Nucleotide
 Genomic coordinates:
 Start: 281865
 Stop: 281127 (SEQ ID NO: 178)

Amino Acid
 MVSTRSMEAKAAAAAKAKEVSPTTSKRKAEDLTEGTEEEESVETHPPSKLPRVDEDEVY
 IDENVGDVQILASSIEVARMERERLAEAMVRDIKIEEKAATEARKEIASRLIYKEMVY
 LLPQLENMTNRLRPRSLRLRHNMETITDRTFSDLQIFNKVTFEFPILTIDIAFLAREKSRVE
 GSRFYNDMKIGPITAYKLNLMCNKFIESVVQVKAEISPFVEVSVSSELEGSPFWDFKQR
 IVKHT
 (SEQ ID NO: 179)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q23915 (Q23915) PROTEIN KINASE | 43 | 0.001 |
| O77819 (O77819) CORNEAL EPITHELIAL RHO-ASSOCIATED-SER/THR K | 43 | 0.002 |
| Q13464 (Q13464) RHO-ASSOCIATED, COILED-COIL CONTAINING PROT | 43 | 0.002 |
| P70336 (P70336) RHO-ASSOCIATED COILED-COIL FORMING KINASE 2 | 40 | 0.017 |
| P70335 (P70335) RHO-ASSOCIATED COILED-COIL FORMING KINASE 1 | 39 | 0.029 |
| Q63644 (Q63644) RHO-ASSOCIATED KINASE BETA | 39 | 0.029 |

Comments:

EST confirmation of the predicted transcript:
 An isolated EST has equence identity to nucleotides 1 to 712 of CT523: this
 corresponds to nucleotides 281143 to 281854 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=404
 5'stop=425
 3'start=502
 3'stop=522
 5'primer=GATCACTTCTCAGGCACAACGA (residues 404 to 425 of SEQ ID NO:178)
 Tm5=58.93
 3'primer=ACGGGCAAGGAAAGCAATATC (residues 502 to 522 of SEQ ID NO:178)
 Tm3=59.06
 probel=AATGACCATTACAGACCG (residues 426 to 443 of SEQ ID NO:178)
 probelstart=426
 probelstop=443
 direction1=Reverse
 Tm1=68.98
 score1=1.98
 length=119

WO 01/38351

139/201

PCT/US00/28888

CT604
 Nucleotide
 Genomic coordinates:
 Start: 50300
 Stop: 49079 (SEQ ID NO: 180)

Amino Acid
 MSHINSTSAATTSSNTLPICTTTAPMIAAARAAIASRTSASAVTSINSNSTSSSAMFRV
 PQGISVTAMPPVPALTSLTESTGTRMSSTPNVDVIPVPGPKNKSKSKKKDSKRKKNQNGN
 RSSDEDEPSLVIDDGSGRQSKNKKYSWVTSLATTTAERNNDTLAPRPFLPTPEEGNMSE
 IDAGLSNPVTRQITGEVYSAALTSGVGDNGLYPHFVADTSYGCETPIPGPAFVLDDG
 TVSRGTSLLHREEAEFLNDGSKVIHTVKPRNSKYSNIQRAASCMAYAVDLLNNHNITSQ
 FDEMAMTAWAARQRCGEMAKFFEKRDKDIGEYRNKVQYNRGIFTRTTEMNKRKIIIEQ
 QQRREAAAAAATGATAPIPTTSAAGVGATSSATTNSLEYQEIRYQ
 (SEQ ID NO: 181)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| AMYH_YEAST (P08640) GLUCOAMYLASE S1/S2 PRECURSOR (EC 3.2.1. | 55 | 8e-07 |
| AAC49609 (AAC49609) GLUCOAMYLASE | 55 | 8e-07 |
| O39781 (O39781) MEMBRANE GLYCOPROTEIN | 53 | 4e-06 |
| O39782 (O39782) MEMBRANE GLYCOPROTEIN | 50 | 3e-05 |
| Q9Y075 (Q9Y075) PROTEOPHOSPHOGLYCAN (FRAGMENT) | 48 | 8e-05 |
| Q14888 (Q14888) MUCIN (FRAGMENT) | 48 | 8e-05 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 666 of CT604: this corresponds to nucleotides 49128 to 49792 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=572
 5'stop=592
 3'start=640
 3'stop=658
 5'primer=GCCAAATCACCGGAGAAGTTT (residues 572 to 592 of SEQ ID NO:180)
 Tm5=59.06
 3'primer=CAGCAACCGTGAAGTGGGA (residues 640 to 658 of SEQ ID NO:180)
 Tm3=59.24
 probel=CTCACTTCTGGAGTTGGA (residues 604 to 621 of SEQ ID NO:180)
 probelstart=604
 probelstop=621
 direction1=Forward
 Tm1=69.03
 score1=1.96
 length=87

CT605
 Nucleotide
 Genomic coordinates:
 Start: 51809
 Stop: 50423 (SEQ ID NO: 184)

Amino Acid
 MDSSASVVFMRFPAPGEETALPPRRATPGSVAYDLFPSEEMDIEPMGLAKISTGYGIDKF
 PDGCYGGIIVSRSGMTWKNNTSVPTGTIDVDYRGELKVILRNHSAEKSVPIRKGTSIAQLI
 FLRYCDVEEEQIVYINETTGERTIIDSSSKKDNKNQARSVRGTGGFGSTDNPNFTETTTS
 RNQREENKKEELEEGEIVEMEGFIDIPFLEGFENILAEQSNETGVITYPNTNQDVEEKDTK
 NIDVVRELEAEFSSGIGSGMDSSDSSSSSSSSDSSDSSDSESSDDSEGGDNKVR
 IRRHQYHRRQLSYSDDVNGGGRNSEKMEMDRVTHIKTEHIKREDEPRYEERERYIHPRRM
 QVPKDYYCEQYEHYDAPAAAHHRHHQHRHQHFNQPRSNSSDVTAYVNENSPTRPC
 RDRNSRFSERPNNGGYNRINSRYTTFDPYRYGARRGRGGVY
 (SEQ ID NO: 185)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q89662 (Q89662) COMPLETE GENOME | 90 | 2e-17 |
| Q86612 (Q86612) ORF2 | 89 | 8e-17 |
| DUT_CHVP1 (O41033) DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEOTIDOH | 88 | 1e-16 |
| DUT_YEAST (P33317) DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEOTIDOH | 87 | 2e-16 |
| DUT_LYCES (P32518) DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEOTIDOH | 83 | 3e-15 |
| DUT_HUMAN (P33316) DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEOTIDOH | 83 | 5e-15 |

Comments:

CT606
 Nucleotide
 Genomic coordinates:
 Start: 65024
 Stop: 64010 (SEQ ID NO: 186)

Amino Acid
 MSSSQGLNNMCTTEILLPKCTSSSLSEESVDYLEKDFEELGIPLVEGKEVLLEFAYKI
 LNKRDITIRVIGDEQGDVCSVFFLRFGKKKTFNPQTKMWLVKLANAIALSMGVVPEPACTC
 SRMMTTAKKIPVPESYKNVNRNIQKFEDVHYIDINFQSFVREQIGLSVLGKNDVQKKKKE
 ETPFFAPFNKSKIGGECIEDLKYDSESVSIIRDVFNLLGEMPTEDVKTSRSCINPSHNDT
 NPSMRLLVFRPMYWRNSKLVMDKLSKEQDSALIEKYMGGHQHCIIGGRNVLLYCITALCF
 SSDCGFKKMLTNDIEIKQLIWYLVLLFFHIICPIIQSK
 (SEQ ID NO: 188)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| YNHE_ECOLI (P77522) HYPOTHETICAL 56.3 KD PROTEIN IN LPP-ARO | 35 | 0.69 |
| TTK_MOUSE (P35761) DUAL SPECIFICITY PROTEIN KINASE TTK (EC | 34 | 1.2 |
| Q56333 (Q56333) FLIL | 34 | 1.2 |
| CAB49508 (CAB49508) HYPOTHETICAL 32.8 KD PROTEIN | 33 | 3.5 |
| BAA85006 (BAA85006) ORF1P | 32 | 6.0 |
| BAA85071 (BAA85071) ORF1S | 32 | 6.0 |

Comments:

EST confirmation of the predicted transcript:
 An isolated EST has equence identity to nucleotides 531 to 1 of CT606: this
 corresponds to nucleotides 64063 to 64593 of the genomic reference sequence.

TagMan Proe Sets:

5'start=546
 5'stop=567
 3'start=655
 3'stop=677
 5'primer=CCCTTTCTTTGCACCCTTTAAT (residues 546 to 567 of SEQ ID NO:186)
 Tm5=57.43
 3'primer=ACATCCTCAGTAGGCATTTACCC (residues 655 to 677 of SEQ ID NO:186)
 Tm3=58.14
 probel=GTATGATTCTGAGTCTGT (residues 606 to 623 of SEQ ID NO:186)
 probelstart=606
 probelstop=623
 direction1=Reverse
 Tm1=68.79
 score1=1.79
 length=132

CT607

Nucleotide

Genomic coordinates:

Start: 68659

Stop: 65032 (SEQ ID NO: 188)

Amino Acid

MGVPEAKKVYENAYGAQNGRVIKEKTGYEDCYDDEDEDYCSGEEDCTTSSLLKATSLAN
 INSKNFLDFGRGKKSSSSSPTCDYTLDMDLPTYNVSDLVMLGRQIATTMLKGQKNMGQM
 ILFINTTNQIIIDVLHDGFNVIREEDTMHSRMQNKKHIYENFYCRDEKKVISEFFSRKYK
 HEKIKARIERVPIIIPSSQEEVDWLTEPPIEDMMMAPPVSNHKMDDYEGLDYWINKHTDV
 MKKRKFLTNSFLFRNVPTTSFNSSPTAVLKSFRKDAFFASQMEGVILYYAFRMIRVMKNL
 LKSKNLKGRYTVLFTDGKAPAIKMMTRAKRQIRQERSKEKAHSRNENCLNRKTNOLLFYS
 CERMMRLPQGLMASALLDIMRIPVLKTTGSKCMYLSNASFTEAEDDIVRLTSCLLNLET
 PGKHFSLLEKRILEYDSYNMSGNRKESKRWEDLLNVLKQHTNDENQTLNLSMNLFSHSDSV
 LVKWNLMVGHKHNVCRLTGTQFKDSEFLKIGHVKFFRCMNSNSSSGENQANELGGFAAKR
 RTKPNTIYNLAESPLMSPESTLLIMLTGSDYNSAIVSNCEYDTWVRKEVAVFENTYCT
 CVGGWEIFLSEQEARKNKDCDDSVGNISMGNLKSKNCRKCDKKLVLPFWTIKFFYLSQA
 IDFVRDPLQLCFPPTHLIDLETDVSLKHALHRAVNAANVMVSYLTMGSEFNQRFVGTIT
 TSLDISIHLSGANNNESKNTGSDVESDTEDLIPFSNNKRKSGNDPQKSTRKSKVNATR
 SAPVTKKLSSSVFESIRGFFESHTEGGIINDRGILTKERIDVFGNNLDTNPEALGEENG
 GGGIVSSIPGLSTEQTSILKTEQNNSTSDFLDFFKKFNEMDDVEEEEEKMEEGEKEEEEA
 DLETDDWLDEARKAFEYKDSDFLEAVTAATNEMTSSLAKNNIEEDEHSRCSVSSKLNNKQ
 PVMDEEKWAEIVNEFDKICSLDNITYNDNSLLSRLSGVLMANKREDGNNSNVVLYEPVQ
 GIDDERFSGVPYSVKTMNLLVIVYMMCGLEDNTIVYQQLMPIIHSEFCGKTEEDKICTD
 RTNFMMSAALEYTMLQYMPKLKTPRIKQIKRKNWERIPKVLDDFKDKVSTCTDNYNKLLA
 TLNKEGKIPSENTKWLP SQGFMPVLGVAISKWSPPLTLWSSFYLQHQQRQDVSLTNITP
 PNSPRPEQ

(SEQ ID NO: 189)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| Q12532 (Q12532) HYPOTHETICAL 119.1 KD PROTEIN YPL009C | 50 | 1e-04 |
| Q9XZ15 (Q9XZ15) RANGAP | 47 | 8e-04 |
| Q13387 (Q13387) HYPOTHETICAL PROTEIN 384D8_2 | 46 | 0.001 |
| Q9YTL7 (Q9YTL7) ORF 48 | 45 | 0.003 |
| Q25662 (Q25662) REPEAT ORGANELLAR PROTEIN | 45 | 0.003 |
| O96266 (O96266) HYPOTHETICAL 283.6 KD PROTEIN | 45 | 0.004 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has enquence identity to nucleotides 1 to 860 of CT607: this
 corresponds to nucleotides 65102 to 65960 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=1428

5'stop=1448

3'start=1516

3'stop=1538

5'primer=CGATTCTGATGTGCTGGTGAA (residues 1428 to 1448 of SEQ ID NO:188)

Tm5=58.31

3'primer=TGGCCAATCTTTAAGAACGTTTC (residues 1516 to 1538 of SEQ ID NO:188)

Tm3=58.35

probel=TGGAACCTGATGGTTGGA (residues 1450 to 1467 of SEQ ID NO:188)

probelstart=1450

probelstop=1467

direction1=Forward

Tm1=69.04

score1=1.95

length=111

WO 01/38351

144/201

PCT/US00/28888

CT608

Nucleotide

Genomic coordinates:

Start: 69118

Stop: 68704 (SEQ ID NO: 190)

Amino Acid

MDGDSSSLQLLSESEFDYVVETLKEQGVWELALEVFNEVSNSIETVKEEEDYTVLRSRNY

FPTESITLYKQQQEEEEESTPIKKRKLASGKSPRSLCRELRLLQIPSTTTFKAAPRSSRR

GKNTRLRRVCKNYGAHQ

(SEQ ID NO: 191)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| VE2_HPV08 (P06422) REGULATORY PROTEIN E2 | 38 | 0.028 |
| FLP_ZYGBI (P13784) RECOMBINASE FLP PROTEIN | 34 | 0.41 |
| Q9XZS0 (Q9XZS0) GUANYLYL CYCLASE (EC 4.6.1.2) | 33 | 0.71 |
| CINA_ELEEL (P02719) SODIUM CHANNEL PROTEIN (NA+ CHANNEL) | 32 | 1.6 |
| GLMS_HELPY (Q26060) GLUCOSAMINE--FRUCTOSE-6-PHOSPHATE AMINO | 32 | 1.6 |
| Q9Z8C4 (Q9Z8C4) TRANSGLYCOLASE/TRANSPEPTIDASE | 32 | 1.6 |

Comments:

TaqMan Primer/Probe Sets:

5'start=114

5'stop=138

3'start=235

3'stop=259

5'primer=GGTGTCCAATTCTATTGAGACTGTG (residues 114 to 138 of SEQ ID NO:190)

Tm5=57.77

3'primer=CGAGTTTCTCTTCTTAATAGGGGT (residues 235 to 259 of SEQ ID NO:190)

Tm3=58.39

probel=AACAACAACAGGAAGAGG (residues 209 to 226 of SEQ ID NO:190)

probelstart=209

probelstop=226

direction1=Forward

Tm1=68.86

score1=1.86

length=146

WO 01/38351

145/201

PCT/US00/28888

CT609
Nucleotide
Genomic coordinates:
Start: 78365
Stop: 77441 (SEQ ID NO: 192)

Amino Acid
MWCSTHLSYSEFFTPLQSQKLFRNFFRALEFRGWTASSTECQVPRVDLWVGPMDSYTRNC
WFQKRTLTFVCFWNRRFWRLVDPEMRGYNLLFSLENFTLPLSQKLFKNFFRALQFRGWTA
SSTECQVPRVDRWVGPMDSYTRNVIAPEYINFCVFLEQAFLETGRPRNERVYPSVFTRE
FYSSSISKTFQKFFRALQFRGWTASSTECQVPRVDLWVGPMDSYTRNVIAPEIEEVSYGH
FWTRCFWTKILLDGNPLPLPPPFFKKGPRVYNDCTTPHSNHHNNHHHHHGRTSILQQTLSRK
WSSLL
(SEQ ID NO: 193)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| DYRK_RAT (Q63470) DUAL-SPECIFICITY TYROSINE-(Y)-PHOSPHORYLA | 37 | 0.22 |
| DYRK_MOUSE (Q61214) DUAL-SPECIFICITY TYROSINE-(Y)-PHOSPHORY | 37 | 0.22 |
| O61543 (O61543) BHLH-PAS TRANSCRIPTION FACTOR SPINELESS | 37 | 0.22 |
| DYRK_HUMAN (Q13627) DUAL-SPECIFICITY TYROSINE-(Y)-PHOSPHORY | 37 | 0.22 |
| HXA1_MOUSE (P09022) HOMEBOX PROTEIN HOX-A1 (HOX-1.6) (HOME | 36 | 0.29 |
| O60275 (O60275) KIAA0522 PROTEIN (FRAGMENT) | 36 | 0.50 |

Comments:

WO 01/38351

146/201

PCT/US00/28888

CT1010
 Nucleotide
 Genomic coordinates:
 Start: 41348
 Stop: 41795 (SEQ ID NO: 194)

Amino Acid
 MTHLVLLILSLSPVYHHLTPYLSPHLYTPISPITSIFPHLIHSLQFQHPVLAEPHN
 QIWTPVFPFIPNRHHLCPQALAVYIRRRGQARSISLQASRRATQALSLLLPRDLPII
 KLQEWPLQPPPHQVLTPCWTLSTYLVRN
 (SEQ ID NO: 195)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|--|-----------------|------------|
| O62081 (O62081) C31A11.6 PROTEIN | 36 | 0.20 |
| Q05330 (Q05330) PUTATIVE ORF | 32 | 2.2 |
| Q60974 (Q60974) RETINOID X RECEPTOR INTERACTING PROTEIN 13 | 32 | 2.9 |
| P79230 (P79230) KAPPA CASEIN (FRAGMENT) | 32 | 2.9 |
| Q9WAL8 (Q9WAL8) POLYPROTEIN | 31 | 3.8 |
| AAD12852 (AAD12852) Y8A9A.2 PROTEIN | 31 | 6.6 |

Comments:

TaqMan Primer/Probe Sets:

5'start=209
 5'stop=227
 3'start=280
 3'stop=300
 5'primer=TCCCTAACCGGCACCATTT (residues 209 to 227 of SEQ ID NO:194)
 Tm5=58.83
 3'primer=GCTGGCTTGCAGAGAACTGAT (residues 280 to 300 of SEQ ID NO:194)
 Tm3=58.29
 probel=TATATAAGGCGGCGCGGC (residues 250 to 267 of SEQ ID NO:194)
 probelstart=250
 probelstop=267
 direction1=Forward
 Tm1=69.00
 score1=1.99
 length=92

WO 01/38351

147/201

PCT/US00/28888

CT1011
 Nucleotide
 Genomic coordinates:
 Start: 41758
 Stop: 42097 (SEQ ID NO: 196)

Amino Acid
 MLDSKLLSSEELKELTSYVSTSSRRSDMKKHLHLFEEHEKIFQFIQGHKFSLYTLDFE
 IFYVMLNILLVEVKNILSPIPLLFDRLNPVRRLLWMFHNGPASPERSRLG
 (SEQ ID NO: 197)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| 044068 (044068) DNA POLYMERASE ALPHA (FRAGMENT) | 31 | 2.3 |
| Q49563 (Q49563) DNA, TRANSPOSON-LIKE ELEMENT ENCODING 3 ORF | 31 | 3.0 |
| BAA74885 (BAA74885) KIAA0862 PROTEIN | 31 | 3.9 |
| 044055 (044055) DNA POLYMERASE ALPHA (FRAGMENT) | 31 | 3.9 |
| 076063 (076063) LEUCINE-RICH REPEAT PROTEIN SHOC-2 | 31 | 3.9 |
| YD89_METJA (Q58784) HYPOTHETICAL PROTEIN MJ1389 | 31 | 3.9 |

Comments:

TaqMan Primer/Probe Sets:

5'start=107
 5'stop=127
 3'start=193
 3'stop=217
 5'primer=TCGAGGAGCACGAGAAGATCT (residues 107 to 127 of SEQ ID NO:196)
 Tm5=57.83
 3'primer=CTTCAACCAACAAAATATTCAGCAT (residues 193 to 217 of SEQ ID NO:196)
 Tm3=57.46
 probel=ACAAGGTAAGCACAGTT (residues 138 to 155 of SEQ ID NO:196)
 probelstart=138
 probelstop=155
 direction1=Forward
 Tm1=69.01
 score1=1.88
 length=111

WO 01/38351

148/201

PCT/US00/28888

CT1012
Nucleotide
Genomic coordinates:
Start: 42053
Stop: 45491 (SEQ ID NO: 198)

Amino Acid

MAPPHLNAAADLLDKVMSGPLSPEGAQTSSPAACVGAKVVKALVSFCQKTRFTTNIVMRE
VKAMEFQGDDFNYSALCASMPQRPVTERQMFALMKSEDEEMGVSANFSPVSDVINPSSL
PSGQEVDSSTSAQISGMFQNVWSLLEECGSGSNSNSPVSRTVLVCTLFIIQVFKFLVTK
VSNVNVNLNQLFGHVVFGLDVAPSNNSVPSTVNNNNKPKSTSNNSNNISNKRVGGSNNS
GGGRSKKVTATAKNPFNNVDGDNHGMFAGAPVDVNLDDFVFPQVETLTSKSTIPKEEVNV
DEDLSKMCRTALTPLIHTFNVFISEINPSKYDRSMFCKGFLTAWDKFVEGDTAGVKRF
RNYILTRSNYASAAARAVYEASIKGTVYYNDKSKFLFHDNVNPDLDKSWGKNKGKKPRIPA
NLMAFMGIDIVKVCAGKIQKYMFAKQFQHPVEEELVPPMAVYAKVAAGLKSGLTFDDDWDL
PEYENCQFIKYDTEGCKKHSELYAKQLRLTGLNQYNKLEEQSAFFPANIVTVTSASSDD
IHGDTIELMYKTKDGVKGVSKI EDENI IKVNPAAEEKNNRVQAEKTLFEIDSDEVCE
RTEEEFFRPTSVVAAPTTPLVPSNVEEEEEEQMEEEEEEVEREEGSDKEDDGDAQAQ
EEMEEEEEEEEQEQQPEEESNGNENQEEEEQEQQQQPEREENKDDSDSDSDSSSSSSSSSS
SSSSSSSSSSSSSSSENEAEKKKEEEVPAKIQKRRLSERPSEAASSPKRMVVEEQQ
QQLSPSLDILQTAVDEMMEEI PAPEPIVATTSPKAATLALKTGFSYSSFVRGDDLSVAGN
TSPTPEAAVPAATCTSDVGNDFLDMLDGLPGDIVMQPGECDVTAKFFEGITLPDGT DNE
CTGFDDLLKATETDNIITTCFTSPIHPSSNSAPRKDIDNCSSIKRSRAGSLFDTDDDSE
TNEVEKEAPKRKKHLKRRNKSHRGSSGASSSHCMSSDEESEDERDMKSTSKVHKSPKA
HVKHSPKYDAVNSDVNNSYNVNSTTCMSSSDSDAEAQPKSHNKSRSRKHSSSSTSDKKQ
NQQCSINTQNVKTVVQSPPSFRSFS PKKDELGLDFLSRKHTKPVVRPYNKKRDVNTTNNV
VQRSA
(SEQ ID NO: 199)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|--|-----------------|------------|
| Q9YTL7 (Q9YTL7) ORF 48 | 101 | 4e-20 |
| O95815 (O95815) DENTIN PHOSPHORYN (FRAGMENT) | 99 | 1e-19 |
| GAR2 SCHPO (P41891) GAR2 PROTEIN | 97 | 6e-19 |
| AAD46501 (AAD46501) LATENT NUCLEAR ANTIGEN | 94 | 6e-18 |
| O40947 (O40947) ORF 73 | 92 | 1e-17 |
| GARP_PLAFF (P13816) GLUTAMIC ACID-RICH PROTEIN PRECURSOR | 91 | 3e-17 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 955 of CT1012: this corresponds to nucleotides 44515 to 45469 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=1345
5'stop=1366
3'start=1389
3'stop=1407
5'primer=CATCCGGAAGTGAAGAAGCTTG (residues 1345 to 1366 of SEQ ID NO:200)
Tm5=59.57
3'primer=CAATCCTGCGGCAACCTTT (residues 1389 to 1407 of SEQ ID NO:200)
Tm3=59.65
probel=TGCCTCCTATGGCTGTAT (residues 1367 to 1384 of SEQ ID NO:200)
probelstart=1367
probelstop=1384
direction1=Reverse
Tm1=68.99
score1=1.99
length=63

CT1013
 Nucleotide
 Genomic coordinates:
 Start: 45950
 Stop: 47825 (SEQ ID NO: 200)

Amino Acid
 MSSTTSPSSLWDDDDDDDEEDEKDVKQEVSNRPPIFSYMETVSFSDNDEDDNKGEEECFG
 SNFDMFGSDSNMPSTSTAPFPPPSTTTPLPTPRSIMDTSDECDEEGAAAASAPSIAASS
 SIPVGISEAELKKMEKKKRKEIKKLKMMKDPLPHLYVGGEPPVAADYKTRANISLYKVD
 PSIDMCGVAPPQFCAELPTPSIDVYTSSYVFPPTPAMHNKKGSKKCQFLKGRKALRKWI
 HENVCMAPPGKRGGVFLAHLEQRFLAHEGDEYKVPRMFVSRVLNKAFFNLIARADTLCS
 MTFYTNLCWIVNGVVVCFDKDDGGIHDASEYATGENFDTVVFHKREEQKTNGSASKRR
 LTPDTSNMGTS TDVQEFQTMGTNTDMQEFQSMGTNTNPIETSSVGVTNPLNPPPRLVI
 TPLTNDVPELDMWLYSPSRGGGNSRMSANTGTSPLSNTPIPTCFTGGANVVVPNGFVPP
 TFPLECEDDDPSIPNSYNEEDKVHFPFYEYMAKYLSPLVPSYNGQTCNVVQEWFKGSF
 SLAKRRGTVPKFCNSHAFNCNMDVCTAMCKWAKTVIRHGQYCNRCIVRRSCTSMLAYH
 YIVCRDASCDVPKCRERVRNDMD
 (SEQ ID NO: 201)

Top Blast Hits

| Sequences producing significant alignments: | | Score (bits) | E Value |
|---|--|-----------------|------------|
| AAC97971 (AAC97971) | FAS-BINDING PROTEIN | 47 | 4e-04 |
| Q19871 (Q19871) | F28D1.6 PROTEIN | 44 | 0.004 |
| NUCL_MESAU (P08199) | NUCLEOLIN (PROTEIN C23) | 43 | 0.006 |
| AAD56625 (AAD56625) | NUCLEOLIN-RELATED PROTEIN NRP | 43 | 0.008 |
| VP41_YEAST (P38959) | VACUOLAR ASSEMBLY PROTEIN VPS41 (VACUOL | 43 | 0.008 |
| O35613 (O35613) | FAS DEATH DOMAIN-ASSOCIATED PROTEIN (DAXX) | 43 | 0.008 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 934 of CT1013: this corresponds to nucleotides 46867 to 47800 of the genomic reference sequence.

CT1016
 Nucleotide
 Genomic coordinates:
 Start: 52093
 Stop: 54913 (SEQ ID NO: 202)

Amino Acid
 MSEPSVYAFIDIKEIENGWEKEFGLLVQPGQKLAFFRDISYDSSKLDCDAFSCIPSDILH
 SDNEKRVGECNFAEHTSVSFPVKNPEGKTLRHFTACGPGCYRRYKQRDPHTGLPVARGVL
 MQDHVDHETGNKMCEYLNQSLVMWAAVPWIRPGDLTEGYNTTHVPGFAFKEDDERDSKRV
 KYENVVISKAYCDFFKQYYDADSGSCYRSGWMKFVHLMFGQYFTNLSYNLANPKPYNLTG
 NTWSDVVSVLTDNPIVDAGAAPSRSEMDEIITKKKFNVFPSEQTSARQKAENIIRSQYGD
 GVEIDPSSVDALMQFVNREGVVGTEKKSDRLMRVADAVMDAAMRLQVMGLDDSQSRRLLL
 KNMIKMSRNNPEYARHFSSSLKLIGVTLAIKRSVFSKGASAKRKETAINNGEQHRRSRWS
 PETVTEEDALLFARENITEDPKHPAPFVDILHSPDINSSIKSGSSSSIWNILSRISSTR
 KLEEKASVVFVKNLVVKVVRQFLDILEGKLFSDGYEWDNIPLMIGVDQILREVIKAASNM
 CARFASSALESSLVTGFIDSASAITSLAVQLAARTFSVFLEESVIEFVVAASLRLAIQA
 FADLATLAASALTIGIVIFVIQVLGLLDLALGLGWYDHIFSPEDLKKQVLVFRREFAK
 AGNVVDVGVAQPVTPEEIVAINVFLQTEENGEEKKEEGARKSKIDFLQKYFHSTPLMGKKS
 KEVYIQEAAQEYLGGRTMNAFGQRIITAADSDTTTTTQEGRRDDETVTKKMRSIILETG
 QTLKDYSSAVNYNASRLDYVGEEWVRNTALKEETRSNTTSDNLFKKTVSLASMAGAFVLV
 GIGVLVASHITLLRFTNIGLAFAFAGLLAFIALMSISYINMNAMGVVNSDAIYRSTALVG
 DIKTDPRRVGMVQRHVGVGAKYNMITDFVSPMLDEIESD
 (SEQ ID NO: 203)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| O67140 (O67140) L-SERYL-TRNA(SER) SELENIUM TRANSFERASE | 41 | 0.048 |
| YD86_SCHPO (Q10411) HYPOTHETICAL 222.8 KD PROTEIN C1F3.06C | 38 | 0.32 |
| Q88606 (Q88606) (CRUCIFER) GENOMIC RNA FOR RNA-DEPENDENT RN | 35 | 2.1 |
| PME3_PHAVU (Q43111) PECTINESTERASE 3 PRECURSOR (EC 3.1.1.11 | 34 | 3.6 |
| O81424 (O81424) T2H3.2 PROTEIN | 34 | 4.7 |
| O04260 (O04260) HYPOTHETICAL 49.5 KD PROTEIN | 34 | 4.7 |

Comments:

CT1017
Nucleotide
Genomic coordinates:
Start: 55054
Stop: 58189 (SEQ ID NO: 204)

Amino Acid
MDSLNTNTVLLVNDRLGNHRTNKPITEQDVENTLNLNSLERASLLKLYSVFIKEMQSYSG
CIPKNKYTNVQEIFEDGLITFEWRDGTKVHRVSPSSPIPLSTKKSPRSSPSPSPSPSI
KEEEFEEEFEDDEEIIYETDENVEDFINGDGEDSEEEEEEDIIVDDEEEENEENGENKYVLA
FSNHLRRQTAAAAAADIEKKDKNHAVSAHDYTLSALQQQQQKLLQQQQQQHQ
QRSSEKVTSTPNKFNKFLPSNGFSEQTELFVCFDVKIAQYNGLVLDILPIVAEYII
NGLGLKCSMETPPVKPCRRKEVKDVWCQPKTSFENDAVEDKHLAFAESPILQRPDPFPI
KKITAYFCLDDSDVIKNPWGSCPLLKSGSNFRVSEYSRHFNEFSGVKNDDDTSNTCFIY
SQKNPNIEIVSKLNIEFEVMMEGIIITHRKDLFETGILSDSSLATAMAFCHPKARVRNVAL
FYFSVYLPFSKTRKETIKSETDKVHIGSDAIFSPSDNPNI SAHQNNNNNNNNNTSVN
IEDRPIRNNNISRKMTITNYQCMACKERCTNNCTNGNYPDRGNQHLSSHVKGEDFFKILN
NSKVDLSLKKLSRVLI PAPPSGNYTSKFCDRSSMCHSFFCRGIEPVSTSFSSDSFEKTKLV
LYGKVVVDVINSYSAIKTSHNNRIRVFFNSEEKDNKTI PSRAESAKNAFKDILVHECNKER
AVSYFEQNKLSKDGHL SNKWWIELNDLNIMFEKHVEDFYKKCSKVND AESLKDIFNDFE
KTCDDKYKTAKRAIIGAQDPSTSTPSKKENGITRIISTLSEFHSKDEATVSALLDKTMLLG
SRTIMSGVRCVIRNNSVFSGFENKNTNNNWELEIRHYVISMGGA AVTKISDEDLEQFTPV
RGAVSVTTAPNOKLPVGAHQTWKDEQTLKTNTKRNSLYDSYNSKRNNRDNNKIKNRSCLK
SDFNWRTPNISIQEFNANKDDVNKKRYAEVVASAAPKSPSPTSSSSSNSNSSSPPLSPLS
PTVKNSNNKPLYIPPHKRMTTAV
(SEQ ID NO: 205)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| 077033 (077033) TRFA | 60 | 7e-08 |
| Q94463 (Q94463) K7 KINESIN-LIKE PROTEIN | 59 | 2e-07 |
| NUCL_CHICK (P15771) NUCLEOLIN (PROTEIN C23) | 57 | 5e-07 |
| AAD46501 (AAD46501) LATENT NUCLEAR ANTIGEN | 57 | 8e-07 |
| O40947 (O40947) ORF 73 | 56 | 1e-06 |
| Q98148 (Q98148) ORF73 HOMOLOG | 55 | 2e-06 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has sequence identity to nucleotides 1 to 215 of CT1017: this corresponds to nucleotides 57667 to 57881 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=1741
5'stop=1762
3'start=1847
3'stop=1864
5'primer=CGTGGTAACCACTTGTAC (residues 1741 to 1762 of SEQ ID NO:204)
Tm5=59.16
3'primer=TTCCAGAGGGAGGAGCGG (residues 1847 to 1864 of SEQ ID NO:204)
Tm3=59.84
probe1=TGAGCAGAGTACTGATTC (residues 1829 to 1846 of SEQ ID NO:204)
probe1start=1829
probe1stop=1846
direction1=Forward
Tm1=69.01
score1=1.98
length=124

CT1018
Nucleotide
Genomic coordinates:
Start: 58947
Stop: 60029 (SEQ ID NO: 206)

Amino Acid
MKICQISSPTLTLSIPILEGVYHVKQLLHLKVHLDVKGVKQLLHLKVRLDVRGAKQNPWRK
NLCLLKKNVKSQKQLPHLKVHLDVKSQKQLPHLKVHLDVRGAKQLPHLKVRLDVKSQKQL
PHLKVHLDVRGAKQLPHLKVRLDVRGAKQNPWRKNLCLLKKNVKSQKQLPHLKVHLDVKG
VKQLLHLKVRLDVRGAKQLPHLKVHLDVRGAKQNPWRKNLCLLKKNVKSQKQLPHLKVLL
DVRGAKQLPHLKVLLDVRGAKQLPHLKVLLDVRGAKQNPWRKNLCLLKKNVKSQKQLPHL
KVLLDVRGAKQLPHLKVHLDVRGAKQQQLCLPLKTISTSFTHLLCLYMEYGKHQNLQV
X
(SEQ ID NO: 207)

Top Blast Hits

| Sequences producing significant alignments: | | Score (bits) | E Value |
|---|---------------------------------|-----------------|------------|
| O41125 (O41125) | A643R PROTEIN | 36 | 0.34 |
| O77336 (O77336) | PFC0425W PROTEIN | 35 | 0.75 |
| Q65683 (Q65683) | 42K TRANSPORT PROTEIN | 32 | 6.5 |
| Q48275 (Q48275) | HYPOTHETICAL PROTEIN (FRAGMENT) | 32 | 6.5 |
| Q89659 (Q89659) | 42K TRANSPORT PROTEIN | 32 | 6.5 |

Comments:

WO 01/38351

153/201

PCT/US00/28888

CT1019
 Nucleotide
 Genomic coordinates:
 Start: 62203
 Stop: 63019 (SEQ ID NO: 208)

Amino Acid
 MDVEFGFFHGLLSKALLPDEKHQPVIRRLCADD SRNKGEDGCCSFCGRRGTGESNTACLE
 QLIDVCSFIGTVSSIGTIINSNLSTSCSRLQKTS DSYAALSHSSFLDVVYPSLKTTEDV
 LPHSLRAIWNKQLPKLYEKTLPPIEEEDIGYKDYVVSIEDDDNVDDGDQEQMIIDEESY
 KTIGEKSTIELIGMYNNNKFGNEFIRIPLRETALHAQSLRYDTEAKFVNHKDSIPLFYEN
 STCTCKERLIDFSE RQLQQLKQDGM DKPTDK
 (SEQ ID NO: 209)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| AAF02799 (AAF02799) F5I10.23 PROTEIN | 37 | 0.14 |
| O23058 (O23058) BAC IG005I10 | 37 | 0.14 |
| OPI1_YEAST (P21957) NEGATIVE REGULATOR OF PHOSPHOLIPID BIOS | 34 | 0.90 |
| Q24042 (Q24042) HYPOTHETICAL 89.2 KD PROTEIN | 34 | 1.2 |
| Q24043 (Q24043) HYPOTHETICAL PROTEIN (FRAGMENT) | 34 | 1.2 |
| O76871 (O76871) EG:100G7.2 PROTEIN | 34 | 1.2 |

Comments:

TaqMan Primer/Probe Sets:

5'start=292
 5'stop=313
 3'start=375
 3'stop=396
 5'primer=GCGGCATTATCCCATTCTAGTT (residues 292 to 313 of SEQ ID NO:208)
 Tm5=58.17
 3'primer=TTGTTTATTCCAAATGGCACGT (residues 375 to 396 of SEQ ID NO:208)
 Tm3=58.03
 probel=CTGAAGACGTATTGCCTC (residues 350 to 367 of SEQ ID NO:208)
 probelstart=350
 probelstop=367
 direction1=Forward
 Tm1=68.97
 score1=1.97
 length=105

CT10000
Nucleotide
Genomic coordinates:
Start: 3143
Stop: 6956 (SEQ ID NO: 210)

Amino Acid

MVLTLSCTTRRVASSKGNFSKEDAVLGNQFPILKKSNNLSIARPPSIESFSASVEKIFRE
WNESGGGEKIFDISQNEEEWMDIISLVESVYEPVFSKSLKPKDLADKTCLTAAAFALASA
VDEKLTILSGSDGSVLQRTTKVMKKDKPKKIAESLLNNEKWTISILLDRKLTAKKLLSRRGA
LKSAERVEVLHRLNKLKEAPLPHHPSLFDNFSGGKTSAVSAGTVIASDMHFKLVEHIFKV
SFRKWGPCGDKTESGEEDEEEEEEEKKHSISRFLQFMNGHNGQHYHRPESASVYFCDY
YDYLAYRNLNPNEYKLSSMHPGTFNMEDLPFRPFAVPSTYKTELEYKRFVQSTNLPQLSFD
YGEFLCYCIFGADWYKHLGDVVDLSLENSSMISFDSQTLSGVYKNTANYKRLGKKRNGIAD
LAVRSMAEFIRTEAHKALTAEMEEMEEEEEEEEAEAMQPEAEVDFLSVPHLRRKIRQAV
SVLNNFVENDLSILVSNFKNVLTDDTVSGTDTDNFGSSGEFEALSSHLFSLRILDEVHIL
RNTDIQRTLFSTHVSLSKSPPSRVRGSNVFNNNAGNISSLQTYGGIEELPENVLVGLS
GGFEDTDMYSGEDVVVVWDGCDGGKVLSTFNCGDNFIQLHEKTAETFKDDTDLVERIRD
VLQTSKGTGNLKKAYSARKNIYAVLRENGIERPGDDFTEKGIALKDKNQPPPPARSAKI
TVEGVKGGFFSGFRDILETRALTTSYSAETFRDLGGQIVKETEGTLTAATVAETSFSEGLAES
LRSDANLGLFESEDAKTVVFKNDTSRSLLEETRALRANNTSFSSFARMGVQVSADLDAE
FAAEMRETYPDAALEQNLKDLKFEETIPESQVKKLKKIDSYLTPERAGKEINDTELS
KATDSVLGKKLGNVAVTVMNNFVGKVTIVVGASVVAGFLGPAVALVHASRGALNVVDHT
SPKGVISYKIVDFSCADRNTGWAKPTKHPFREEIDHVIALDASFLTENGAYVFPEDGGPK
SKYKAYAPICGTDAAQGECSWATFDDPHSVLPWVASMKDLPGQSLSCDKGMSTLKAV
SSVLLSIGKDVAEAI FEVAEDAVVGLASKAISAVINNPLFIFGVPLGFGIAATRLNPSNW
KTGLIVFSILLVILIVRFFAGSGPLTLNWFAGKNSAKRKQTEQFEDGGGNRSKIVLAEK
DNANSKLQSRNETGPMRLEELPGHEDLRPVFFPATTNYSKSAKILGYKSKPFNDFYTKI
INTDIIKMDR
(SEQ ID NO: 211)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| YK06_YEAST (P36062) HYPOTHETICAL 75.5 KD PROTEIN IN SDH1-CI | 43 | 0.010 |
| Q12532 (Q12532) HYPOTHETICAL 119.1 KD PROTEIN YPL009C | 43 | 0.017 |
| O75184 (O75184) KIAA0702 PROTEIN | 41 | 0.050 |
| O60721 (O60721) RETINAL ROD NA-CA+K EXCHANGER SPLICE VARIAN | 41 | 0.050 |
| O43485 (O43485) RETINAL ROD NA+/CA+, K+ EXCHANGER | 41 | 0.050 |
| AAD28522 (AAD28522) FLAGELLIN (FRAGMENT) | 41 | 0.050 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 969 of CT10000: this corresponds to nucleotides 5867 to 6835 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=2102
5'stop=2125
3'start=2185
3'stop=2204
5'primer=GGATTGCTCTCAAGGATAAAACAA (residues 2102 to 2115 of SEQ ID NO:210)
Tm5=57.56
3'primer=ATGTCACGAAAACCGCTGAA (residues 2185 to 2204 of SEQ ID NO:210)
Tm3=57.81
probel=AAGTGCCAAGATAACGGT (residues 2148 to 2165 of SEQ ID NO:210)
probelstart=2148
probelstop=2165
direction1=Forward
Tm1=68.99
score1=1.99
length=103

CT10001
 Nucleotide
 Genomic coordinates:
 Start: 17098
 Stop: 20842 (SEQ ID NO: 212)

Amino Acid
 MGAPTNADFTRTVSGVASSLYLVNPGAPSDREKLV LASSYSDSFVYNYKDAVVTAEPKW
 CPFNEPALHEHIMNRLEKAGLINRSRVCNPVKSAGEVCGFRYSGGSTPQNLIFPIGASE
 INWLVLIRNAARFGTVAAASAKDAIERIPDLREGGTSKHVAKNAMRRLRVWRAFNWIAEAS
 RSAGMIRYEPFVSGCALYDHEVRRRQLKGSYERNVFLGDHFCKESTLLADISRGRSSD
 FWTIVEAVIRYKNRHARTISNETNAIPEDSSINLEWEDVLVKNQORDTNQGDDSTLEKTL
 AAIKEHESIGEKRKKHILEFIKTCLTEEQREMIFKGVGGKGNLSPAHLTNLADAILANNA
 KAGIWWIMQSLLKQINFSLHLIGYEAQRLLMFKLYMPALLALFISQRGIGDVFLNGVFN
 LEVMKERAANAKIRDVMSRDAYKNNNTNESNNLGNYSQFDIATIIYQMSDCNAIPLSINIG
 VPLHKSRMNMQDIEKTIQHIVDMGLAKIRVAEENRVARRLLHRRKMQEKAARERAAARQR
 RLRGEEEEEEEEEEEEEEEEAGTSGVNGSGYDQEEEDDGEDEEEEEEEEEDEEDS
 EGEDMNGENNSRKRKNTGNTSSTQPPQKRQRGKNAPISTKGKRLKERDNIGGFLLATIQ
 NDDRQVNVESIQLLTARQRKYVKGKCGDNLPELLIEKVTNLLDSVVFHFRKGSILNSI
 HANRRSETGVYTTKANCICDYERNVSKDSSNNTPHSSECIERARERDASCAESNKRPCP
 VDSNNPEDVEORMRELIMDPPSLSGVEDSLAIERVLQNEILFTSLVTNPINFNAVLGAEKG
 DLGRFIVLNNIVKFMNMTIACLVGDMPMLLDSRGKTKNLEKGTVKNTRKFFKPNMTAA
 ELNVATAQSAGHQYMNAGHCPEPGIKQMLLPDCIMKLKSIAMEKGRGGRSALHRQKCDHA
 FCKMLKCLFFNIDPSNAADTFIDPASRATLFRLLDLCRDRKKYKNIDWVKDLLDPVMKGT
 NKWVGTGEYTNIGRDSNVAAPVDFYTLKYTMIDDGVISVPSRKPNDDVYYSTIERADDLL
 TESRDASCESYRPTLFDARAVLEVNGDGRVPYPSSSETVEDLGEEEEEEITGVIDDSTEI
 EDVQVQDSNLFDELFDIPEIEQHQQGEEEEELPSAISEVFASLPADNDSSSPAHIPTFG
 NSEEGEKSPEPYNIFDSALDQLLDLIDSDGRNNNPKRVDWNSVTIQE
 (SEQ ID NO: 213)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| GARP_PLAFF (P13816) GLUTAMIC ACID-RICH PROTEIN PRECURSOR | 77 | 6e-13 |
| O94922 (O94922) KIAA0835 PROTEIN | 76 | 2e-12 |
| Q9YTL7 (Q9YTL7) ORF 48 | 75 | 3e-12 |
| CENB_CRIGR (P48988) MAJOR CENTROMERE AUTOANTIGEN B (CENTROM | 75 | 4e-12 |
| Q07034 (Q07034) RNA BINDING PROTEIN | 73 | 2e-11 |
| NAB3_YEAST (P38996) NUCLEAR POLYADENYLATED RNA-BINDING PROT | 73 | 2e-11 |

Comments:

WO 01/38351

156/201

PCT/US00/28888

CT10002
Nucleotide
Genomic coordinates:
Start: 20783
Stop: 23729 (SEQ ID NO: 214)

Amino Acid

MEETITLKESTGTVPFKNENITRIASNYVRAFTDTWSHLVNISGAPLTAENPSAIPAN
ELNRYWTKTNVLCNPLFKLEDHITRDEDTGTITLKFKMYIDKNGLYQSAVLMALDSFV
SLASFSGADLVSNKSENKFCVKIPHDTRAESLLNNVGFPAGLSGPFKRSINYKAANLS
GKSGIDGLSGSMLTVLKNNTNKRATDILHLVNNVSASAAQLDDSEMSRTFNHQKKGVCY
DINVSSSRQVNQRNLLHHQNIIGQHLIEFRTKQLERAQNKVKKEEENGEHEEMTSEEEEE
EDEYEEGGCLSDIDEEDFYEDGYDEEEGDDNRTRKKKKMEEDEDEEEYDDEEDEEEAE
TCGANGVIDCEDDAIIFPNGQNSKRKNGKKTNIKKRSRRKGECSANTLSFVEKYVGNCK
SLGIKPVGCPPPSTEFTSLFMKGSEADSCYNTCQSTRGASRIRSLLNKYSVKDLMQVNSP
SSWKWANPPDRRFVLFDKKTKEEVEVFEIECEKSEYFDVSELPSNIKVWLKETAKIIK
HLALIEDFLPAMGAATPKIPLNLIKTMTSIFSVRDIVGFKIPEEVLSFPIPIEWKTSISAM
GLLSVQFDRIIEVIDLMTNGAFATSCLNNAFFLERGVVPRDGSNTWLHTDLVQLSTSIF
RSIRNRGVNIGGNNTGSNSSSSSCGNGKGDYGVRCLGISKRGITLKPPPAAMTNSSSP
SSSAMISLPQPTRQSIDLSITTIIQDFSEVSGKRLNLGLQKNMSDKSKDVFNDIAIYDSGA
FKALLTCTVNDKSRKRKRRTLLASGEGVRRNLMVSQGNVDNDAHQFQEECGIKIGGGA
SRVYKRAQRGSASVSSRRVRNKPQFTIAVSEDEDDCEEGDFSSSELNPTHSQLLLFQQR
QQDSCTEDDDVLVSVEEYNNRVSGSSTTAGDRVLAKDLLSTVSPNEKRNSAALAALTISR
HSLFNALSAKTKLGENGRFFL
(SEQ ID NO: 215)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q9YTL7 (Q9YTL7) ORF 48 | 75 | 2e-12 |
| GARP_PLAFF (P13816) GLUTAMIC ACID-RICH PROTEIN PRECURSOR | 75 | 3e-12 |
| Q07034 (Q07034) RNA BINDING PROTEIN | 74 | 4e-12 |
| NAB3_YEAST (P38996) NUCLEAR POLYADENYLATED RNA-BINDING PROT | 74 | 4e-12 |
| Q9YPA9 (Q9YPA9) HYPOTHETICAL 45.2 KD PROTEIN | 72 | 2e-11 |
| O40947 (O40947) ORF 73 | 69 | 1e-10 |

Comments:

TaqMan Primer/Probe Sets:

5'start=1699
5'stop=1723
3'start=1773
3'stop=1797
5'primer=ATGACGAGCATTCTCTGTAGAG (residues 1699 to 1723 of SEQ ID NO:214)
Tm5=57.07
3'primer=TGCAGAAATAGATGTCTTCCATTCT (residues 1773 to 1797 of SEQ ID NO:214)
Tm3=57.23
probe1=ACCAGAAGAAGTGCTCAG (residues 1743 to 1760 of SEQ ID NO:214)
probe1start=1743
probe1stop=1760
direction1=Reverse
Tm1=68.95
score1=1.95
length=99

WO 01/38351

157/201

PCT/US00/28888

CT610
 Nucleotide
 Genomic coordinates:
 Start: 85707
 Stop: 83427 (SEQ ID NO: 216)

Amino Acid

MPPKHKPNTALKKHIIRNQQRKKEDDAESRFQRMGQEVSKLDAPTSSKNRQRRKIRTSK
 ILSRSGDCVAGDCSDLNDEGKRDTDQEGGGRGGGNEEEEGKEEGEGEEQQREEKEEQS
 EEKEEKDGEEEEENVEDEHVTPTTSVSKRAKQMKKHIFPPSKKRKRSDESKALAVPAG
 KMMTVSRPLRGAITSGSILGVRSENAPODYVSYLADEAVVKEKAIQYRIRSLANLLKA
 NKTKAFPTSSSLLSSEQGKKKFGGKRTNTFVVTNVGAELVKALLANSCWAIHRKDIRSG
 EIQWQELSSKILKSLNDGNATEINNLMSSIVEDRIQRTVKERVYFEQLATVCNNLFGTRI
 LPNKNFDKNFVSVASDNSNATVRGLSIPRYFRAINNNVWVKMSSTMDLLVGGGMRSEH
 SISMLEKCAAGVLARASARPVEKMIKSAVEETSQAFNLSTGVFVVPKQQQQQRRQQQQQFP
 PFQPPPFPLPPPQAFQVQPTYQGYLNPYYQYNQYYPYAPQQLQQQYPLYFLGNQSQP
 PPQLQQQQQQQPPQPPNNIPPPPTPQQQSPSNIPPPQQQQQPPFVQLISSPPPPPIIP
 NTAPSPPISRVRFDSRSTTPQPPPTPVLPKPTPLPPPSTARAEENATDMSFTDIDSELG
 SIDFDLPPATPGRNVEEIIKAQRQAVKETGVRGEEEEEEAFIPIIRQPRTPGNFRDEL
 LDVNESIYGSIDIEPAAAAAFDWDMLDLDLNGDEPYEFE
 (SEQ ID NO: 217)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| P93794 (P93794) LOW-MOLECULAR-WEIGHT GLUTENIN STORAGE PROTE | 84 | 3e-15 |
| Q05573 (Q05573) OMEGA SECALIN PRECURSOR | 83 | 5e-15 |
| AAD43602 (AAD43602) T3P18.1 | 83 | 8e-15 |
| O04365 (O04365) OMEGA SECALIN | 83 | 8e-15 |
| O48809 (O48809) F2401.18 | 83 | 8e-15 |
| O65375 (O65375) F12F1.9 PROTEIN | 82 | 1e-14 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 34 to 361 of CT610: this corresponds to nucleotides 83630 to 83957 of the genomic reference sequence.

WO 01/38351

158/201

PCT/US00/28888

CT611
 Nucleotide
 Genomic coordinates:
 Start: 88938
 Stop: 85761 (SEQ ID NO: 218)

Amino Acid
 MEGVTSIVAAVPEVAILITDLMGGRNNKRSTYERIVGIVGESGDLEAILDICNRNSYR
 DELLEGETVVINPTGLLKEISLLMKKALDMNIKMSNDDPVPFTTLDQNEQEFIGHLKSC
 KKQDGPAYKDLIHRISGMFVMKNTRLMLDEIIRGNAGDAVEEKNALCEAYAEMISDMDL
 IRIFLLLVAIKRDQNKKHRHMKSVIYEDVVVSLNTLKDVFHKEWYMWPFSAQVGTKIRD
 ARTFSVLFGSDMHEGRNNDRIWENMAFSVTEAFLSGPSTNNHYNKGHLRMYAARPVYDAM
 EYVPQELHHILFGTKIAKMIDIVYRYSIYNVPYLLAADTERVEEPKKSVMSPSGLIISP
 ASLLENTPLSLVSRHGIPSAKLGSGFILEHENAENMHLEAIAKCMVSQTLQEESWGESQ
 AAMVYQPSDEVEVIQAHVTKILSGNTTNKTCGLCYADLDMKPKFFNCSENMKASYDYFP
 VHAEMDTFEARQETCSAKLCPDCTIKHLMYVYEKVSAGSEKLDVFRCPCCGEYMWQFIG
 RCHFSSSLFERAILAGENVDPYIAANKLLITELIKRAEKCFTVELLQAEFMEMCKMDK
 DFALDKDSKFTVVDNRFRPPVKLFKMVEGETGDSKCSLICTQCLLPNVCDQPNEMEDIVT
 VDVEPPVLPYPPEQLEDYFQDVEDAEFDDPPTDELVRDYGPGHLKWPMLSCGFLAS
 NFVPPNEEVTNCRQAVSILKRTPEKKIRGWNPEPESPEGVLLALANWHSTDRMPENMKGLL
 NDISVIHNTRETFQNRVKVHYLNSVFGGFDDRDFEQVGVSIPIATYFYVEKLNHESA
 LGLWAKMFVKNLIGEMVLERPECVFRHRAHSFVLHCVDRRALSGIRPNQGAKEIVKQVNI
 VRQNMTSSEIKDPVFTVDEKRTLEWKVEKEGQEIKTVCCKPCKTPNLIKGGCITMTCYDC
 SGRRDGYPTVFCWICEDEITNPDHILIDHKLLYSCKSTKAALEKVYNCTLCCLALARKCS
 DSYLSKQRGSGEEIEIYVMEDGFEFDVHTKTAVPTK
 (SEQ ID NO: 219)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| AAF04637 (AAF04637) HYPOTHETICAL 84.4 KD PROTEIN | 1495 | 0.0 |
| BAA78677 (BAA78677) HFB30 PROTEIN | 44 | 0.006 |
| AAD21842 (AAD21842) ANDROGEN RECEPTOR ASSOCIATED PROTEIN 54 | 44 | 0.006 |
| O94793 (O94793) HRIHFB2038 PROTEIN (FRAGMENT) | 44 | 0.006 |
| Q9XII0 (Q9XII0) F7H1.11 PROTEIN | 41 | 0.042 |
| YKZ7_YEAST (P36113) HYPOTHETICAL 63.6 KD PROTEIN IN YPT52-G | 41 | 0.056 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has sequence identity to nucleotides 1 to 829 of CT611: this corresponds to nucleotides 85988 to 86816 of the genomic reference sequence.

Hit to public SBV sequence:

gi|6165655|gb|AF099142.1: CT nucleotides 1 to 2738 match nucleotides 8582 to 11319 of the public sequence with a 99% homology, a score of 5398 and an Evalue of 0.0

WO 01/38351

159/201

PCT/US00/28888

CT612
Nucleotide
Genomic coordinates:
Start: 91607
Stop: 89060 (SEQ ID NO: 220)

Amino Acid
MGSNQQSFISKRNQKQEIISLEKIIKRIENACLPVNQYVPKLDKNAINPQELASHIMDR
LPATISFQEMDDFLADYAKTKIVDHPDFGKLAGRFICSNIHKNTKEWNSFSATTQKLRHA
IHPGTGKPASVVNDTYEYENVMANAEILDVIDYKMDYLFCTCFGLRTLEYSYLIKIGSPTD
RKKRILVERPQDMIMRVAVGIHGSIDIKSVIETYDLMSRHYFTHASPTLFNCGTVTPQLSS
CFLGLQDDSIIEGIYDTLKEAAIISKTAGGLGIHFHDLRAKGSPISSWSGTHPGLMAFLQ
IFNVSVKKVSQGGDKRRGAAAIYISDWHLVDKDFIDCRKNAGNEDLRTDLPFAIIVVSDL
FMERVKAGKNWSLMCPHECPGLSDVHGEEFKALYKEYEAEKGKKEVVKARALFDQINSAR
IETGTPYVCFKDTINRKSNOENVGIIKSSNLCTEIVQYSDSEETAVCNLASIAVNKFVKY
SPIPSLRPYVDYREMKRVVKIMTRNLDKVIDVNFYAVDKTRISNMKTRPMGLGVQGLADL
FFKLRIPESEEAALINKRIFETIYYGALEASCEIAKEGETYELFEGSPLSKGIFQFDM
GKENIKNRDIYFNSLPIDHWEQLRRDIMKYGVHNSMFVAPMPTASTAQILGNSESFEPLT
SNMYNRNVLSGSFQVVEYVIRELIKLGWNSVTKQRIMASGGSIQTLNIPKSTKELFK
TVWEINPRTTLDMAIQRMFVDQAQSLNLFVEEPELSKVRSMYAWEKGIKTLYYLRTK
GAARAVQFTVDKNVLQEVKKEAPSPVAAFSAPVREEEEEKSSIVVDPAAALLCSINNP
GACEMCSS
(SEQ ID NO: 221)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| AAF04636 (AAF04636) LARGE SUBUNIT OF RIBONUCLEOTIDE REDUCTA | 1720 | 0.0 |
| RIR1_SCHPO (P36602) RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE LA | 807 | 0.0 |
| RIR1_MOUSE (P07742) RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE M1 | 804 | 0.0 |
| RIR1_HUMAN (P23921) RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE M1 | 803 | 0.0 |
| AAD37491 (AAD37491) RIBONUCLEOTIDE REDUCTASE M1 SUBUNIT | 801 | 0.0 |
| RIR1_YEAST (P21524) RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE LA | 796 | 0.0 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 845 of CT612: this corresponds to nucleotides 89225 to 90069 of the genomic reference sequence.

Hit to public SBV sequence:

gi16165651gb|AF099142.1: CT nucleotides 1 to 2547 match nucleotides 5913 to 8459 of the public sequence with a 100% homology, a score of 5049 and an Evalue of 0.0

TaqMan Primer/Probe Sets:

5'start=1327
5'stop=1348
3'start=1383
3'stop=1402
5'primer=GTCGGCATCATCAAGTCTTCAA (residues 1327 to 1348 of SEQ ID NO:220)
Tm5=58.78
3'primer=TGCACACTGCAGTTTCCTCC (residues 1383 to 1402 of SEQ ID NO:220)
Tm3=58.45
probel=TGTCCAGTACAGTGATTC (residues 1365 to 1382 of SEQ ID NO:220)
probelstart=1365
probelStop=1382
direction1=Forward
Tm1=69.00
score1=1.99
length=76

WO 01/38351

160/201

PCT/US00/28888

CT613

Nucleotide

Genomic coordinates:

Start: 94397

Stop: 94175 (SEQ ID NO: 222)

Amino Acid

MVIRLCFLESITCFVYGIMAPLSLDTNTDYLSHKKDTNKKIQMQINFIPYSNMHVYIAGV

YTFHEKKGLTYQQY

(SEQ ID NO: 223)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|--|-----------------|------------|
| O96794 (O96794) AMINOPEPTIDASE P | 29 | 4.1 |
| YA2G_SCHPO (Q09706) HYPOTHETICAL 157.7 KD PROTEIN C2F7.16C | 29 | 4.1 |
| O24567 (O24567) ESR3G2 | 29 | 5.3 |
| AAC36183 (AAC36183) PUTATIVE PEROXIDASE | 28 | 9.1 |

Comments:

Hit to public SBV sequence:

gi|6165655|gb|AF099142.1: CT nucleotides 1 to 222 match nucleotides 3447 to 3668 of the public sequence with a 93% homology, a score of 357 and an Evalule of 1e-101

TaqMan Primer/Probe Sets:

5'start=48

5'stop=65

3'start=169

3'stop=194

5'primer=TGGCATCATGGCACCCT (residues 48 to 65 of SEQ ID NO:222)

Tm5=57.92

3'primer=TCATGAAATGTGTACACACCTGCTAT (residues 169 to 194 of SEQ ID NO:222)

Tm3=58.55

probel=CTTTGGACACAAATACCGA (residues 68 to 86 of SEQ ID NO:222)

probelstart=68

probelstop=86

direction1=Reverse

Tm1=68.82

score1=1.22

length=147

WO 01/38351

161/201

PCT/US00/28888

CT614

Nucleotide

Genomic coordinates:

Start: 129006

Stop: 127158 (SEQ ID NO: 224)

Amino Acid

MEYIGKNNNPVSNESVSEKELKLRSSFLMIGKKTSKYEQVMGVYEAIESIRQSESEDT
 FVVHVKKDKQLKFARGLKRLQELVEDDSLRIERISCAPPEPGHLFKDDAGHVTDEEWLAT
 QEEDVRKINTIVKEKLKRKDKDFKFSQLYRYMSNSLSEAVEKKHDCMISSDFLIGLGFS
 TMNVTHALKSMERTMQKHGFKDMMVPLVEICHRTHYKGEYIANPIFKSHSSHCLIVPLFM
 VAGVFARSAHPAASIEMYLSLAYAVILYSEKQROIREELARKNLQIKEELENQVEKT
 TKVEKELETQVVKTTKVEKELETQVVKKEEYKNSYIETEQLFKVSEEQKESLRNVHKKSS
 NATFRYDSGSCLVFSISSTEFYLLCRTDKSGSFETATENGLRYIFSPINKKRDTAGMRPR
 LIMAVTGCDAPIACNDSIKHQKFKVLKCNRSSIVFQTPPSDEDLKGIVQKVTGSDIRIF
 MNDGTVYQDGGRIDISSPQELDEENMTQFEIEQQRKLHSMMENTS KIVTRYNKERHLTTK
 EARTRNKTEKWFEKVKKREEQKKRENGEQSTSEQEQRGVKRTWENDNEFDSVVEEEDGN
 NTQEQQRVKRHAISV
 (SEQ ID NO: 225)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q06166 (Q06166) MATURE PARASITE-INFECTED ERYTHROCYTE SURFAC | 54 | 2e-06 |
| O66878 (O66878) CHROMOSOME ASSEMBLY PROTEIN HOMOLOG | 53 | 4e-06 |
| GLE1_YEAST (Q12315) RNA EXPORT FACTOR GLE1 | 49 | 8e-05 |
| O35788 (O35788) CYCLIC NUCLEOTIDE-GATED CHANNEL BETA SUBUNI | 48 | 2e-04 |
| Q14789 (Q14789) GIANTIN (GCP372) (MACROGOLGIN) (GOLGI AUTOA | 48 | 2e-04 |
| Q9ZU69 (Q9ZU69) PUTATIVE VICILIN STORAGE PROTEIN (GLOBULIN- | 47 | 3e-04 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 2 to 976 of CT614: this
 corresponds to nucleotides 127217 to 128191 of the genomic reference sequence.

WO 01/38351

162/201

PCT/US00/28888

CT615

Nucleotide

Genomic coordinates:

Start: 130290

Stop: 129405 (SEQ ID NO: 226)

Amino Acid

MEYIGEQLINLLDETPEEDELQLRSSFLMIGEKQYKEYEEVMSTFEAVETIRKSEFRDG
 VFIVQLKENKHITFEGGLKELRELTGDNLSLKIESLLSSIKPEKGVILKNTSTTTDDEWL
 ASQDKDVQEVNKLVEKTRMLFRGFYFSPTYRYITKSLPQIPFGEKERFVVSTDFLIGLG
 FSADDVMEKLIAIEGNMRKSGLKYTWPVAEVCHLKKYKGDIVVNPIFKSYHSHCLVIPL
 VYLGVMFSRNVQPPSLEVETYLLALAFADLYGREEMRKSCMLCEDISEVKRG
 (SEQ ID NO: 227)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q13779 (Q13779) APOLIPOPROTEIN B-48 (FRAGMENT) | 36 | 0.34 |
| APB_HUMAN (P04114) APOLIPOPROTEIN B-100 PRECURSOR (APO B-10 | 36 | 0.34 |
| Q13788 (Q13788) APOLIPOPROTEIN B-100 (FRAGMENT) | 36 | 0.34 |
| P96470 (P96470) IGA-SPECIFIC METALLOENDOPEPTIDASE PRECURSOR | 36 | 0.44 |
| O28789 (O28789) SIGNAL-TRANSDUCING HISTIDINE KINASE | 35 | 0.58 |
| CAB55172 (CAB55172) HYPOTHETICAL 77.9 KD PROTEIN | 35 | 0.76 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 725 of CT615: this
 corresponds to nucleotides 129499 to 130223 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=296

5'stop=318

3'start=360

3'stop=383

5'primer=TTAAGCCTGAGAAAGGACACGTT (residues 296 to 318 of SEQ ID NO:226)

Tm5=58.14

3'primer=TGTACGTCTTTGTCTTGAGAAGCA (residues 360 to 383 of SEQ ID NO:226)

Tm3=57.87

probel=CTACTGATGACGAGTGGC (residues 341 to 358 of SEQ ID NO:226)

probelstart=341

probelstop=358

direction1=Forward

Tm1=68.84

score1=1.84

length=88

WO 01/38351

164/201

PCT/US00/28888

CT617

Nucleotide

Genomic coordinates:

Start: 148612

Stop: 147766 (SEQ ID NO: 230)

Amino Acid

MSPVISQSSPSATSTAAARIISTANLRVLGVKNKEEKDEEEEQQEVEPEIIEPATDFEI
 PFSPALTICIIYINANRIHINSKGVCLNRKKIKPTSTINKNQDVPPELANASSYLVQTEHV
 TDKFLSSHCSICNYNVNDGEYKSALSTTRNGDQPLMRKSVRYVPLNEDNVVVQKGTYYGT
 TFIPEKTGRRILWFESHYKKSPPITAKLCCLLETINSFNGSCSSSSSSASSSSNAPGPIEE
 FQVSSSIFFKKECCPLQMKWVEQNELDAESPVLVLLMLAL
 (SEQ ID NO: 231)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q9Z2U2 (Q9Z2U2) ZN-15 TRANSCRIPTION FACTOR | 36 | 0.24 |
| Q20150 (Q20150) F38B7.5 PROTEIN | 35 | 0.70 |
| O94002 (O94002) SEC12 HOMOLOGUE | 34 | 1.2 |
| Q20497 (Q20497) F47A4.2 PROTEIN | 34 | 1.6 |
| O13779 (O13779) HYPOTHETICAL 71.2 KD MEMBRANE PROTEIN C17G6 | 33 | 2.1 |
| P93002 (P93002) REGULATORY PROTEIN NPRI | 32 | 3.5 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 741 of CT617: this
 corresponds to nucleotides 147819 to 148559 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=363

5'stop=383

3'start=446

3'stop=467

5'primer=CGACAAGTTCCTTTTCATCCCA (residues 363 to 383 of SEQ ID NO:230)

Tm5=58.62

3'primer=ATCAAGGGCTGATCTCCATTTTC (residues 446 to 467 of SEQ ID NO:230)

Tm3=58.15

probel=CGGCTCTAAGCACAA (residues 428 to 445 of SEQ ID NO:230)

probelstart=428

probelstop=445

direction1=Reverse

Tm1=68.92

score1=1.92

length=105

CT618
Nucleotide
Genomic coordinates:
Start: 150145
Stop: 148675 (SEQ ID NO: 232)

Amino Acid
MESVRDVKFYTFMNVLAEKAKKIQRLNKDKGWRTSINAEIGYGGARLMDVRFTRGRKSMDE
LARCLYNCDGEYTTLRVLVGSSAGNIIVYSLAFIMGIRGECCGFNVNNRLRMGKIIDRELF
YKITGLNFPETVKCTCDGVRAICDLFLEVAALQEHFAWHETKEVGKKQQQHFNFGSQYP
GTKFNKRHKLSTKIIQQMFSEKTEQVLAFSEGTAASGFSPLYVEAPIQYVVNMYRAIS
NMEGRVGMYNLSRVLILLCSRWEKKPGYKNDYFYSKCEMYIGSKKIVDDSFIFDLITG
DLVPLVRLAPSNEIDQRDVIRFNDSTDIILMDSIDVRDVLPVLSKIIWQNVSARLKLNN
KSLSKLAKWKWNGMVSTHDFDSNDYVIEHQRQLAADIMSDSLSKNHLPNFSKTITEYDE
KENKTTPLICWNYIFELSPMGKHLFPLEEVCGFYEASLPLITPWQLKVVQKKRGRQMVY
GPRKRPRTQ
(SEQ ID NO: 233)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q64568 (Q64568) ATPASE (PLASMA MEMBRANE CALCIUM ATPASE) | 37 | 0.28 |
| ATC3 HUMAN (Q16720) CALCIUM-TRANSPORTING ATPASE PLASMA MEMB | 33 | 5.4 |
| Q12019 (Q12019) HYPOTHETICAL 559.3 KD PROTEIN | 32 | 7.0 |

Comments:

TaqMan Primer/Probe Sets:

5'start=676
5'stop=701
3'start=823
3'stop=849
5'primer=GAAGCACCTATACAATACGTGGTTAA (residues 676 to 701 of SEQ ID NO:233)
Tm5=56.76
3'primer=GGAGCCTATGTACATCTCACACTTACT (residues 823 to 849 of SEQ ID NO:233)
Tm3=57.26
probel=AGAGTAGGTGCCATGTAT (residues 733 to 750 of SEQ ID NO:233)
probelstart=733
probelstop=750
direction1=Forward
Tm1=69.06
score1=1.93
length=174

CT619

Nucleotide

Genomic coordinates:

Start: 169814

Stop: 165116 (SEQ ID NO: 234)

Amino Acid

MSTTQTQTIERPLPGKNNEDNSRLACLLAEGLOQQQQQQDGDSEISLPLVNAGTFACYDS
 TLANLTEGRLGSETENAKIRVKIHPSVFIETNKEMTIEEISTKSLNALVEKRAREARRF
 SSLTEQKFPRGGGGCYSRKNERFIEGEINNIKLNMEETASSLERLAGLLPVVINIKDWTM
 HDEKEIRLDLKGNDGMEELVNISHLNQEEWEMERLSSSIVLKDAYGVFYAHHGILDIVLT
 TSRFTGKLLQHPVIFRLMDVKVWINTPLQIAFPDTSKNPNNAKKILYQHPSLTRLRDLNDM
 ASNSKSVSSIIPELSKFNSTEFGMHYFTAQCFFGKNTNSLKDVLTRYQYLSFKNKPQPK
 LYEPTATATAASSSSSTASLTTEQKEKIAQSILSSKGKSLGDSSTLSKEYDENRKRTRK
 QKTSTDNTIVPSGAPTSISMKNPVTCFFGPQYTSIMDCISEKTDWIEMHLFLTSLNDAEH
 NKTIVVDRKSNVSEIHDSGRFLTFGQNNNTAFIPDVIDPTLKLILRDDSCESSAIIASL
 IYYNNVNLEGREFNVSDAVVGLFSGGSAITVGDIAREIASIYNIGRESNCDSILFPGEF
 ILAGRRSYGRQYRWYDPINCVVGLYRSCLETMTRNIMRGQPVKVDATAWMMHQQVLQVV
 LLPFFDCVLKSGVWAVKEARQLTDYIVREVLLKYTADPDQHKFLLEKFPVMDLIAKIVTH
 YAVIHSAADNGGVCIAFPDPFPFIVENDTSLRYTTLTDPQSILNGDNVAENLKSATSVA
 SSPSSSSRYSSSETPIRVVNLVPVPTGRFLKMNKDLELFINVPLISSKEQKQQQQQTATAP
 FSSETISKSLNYPVKSLTRNVTYGQNIADGFLGLKNKGELVSYFKVVKNTERRDDGIK
 DMEIGDINNHDQDGLSSSSSSSFDGVRTSFSVDGKIEHVSAFLPGTTSQPTNLPVHAS
 KQVKYSVKELGMETVFFPEPLSSAVLYEASKTKSTQHLSPMRIYKECVSPLSTGRIDIFP
 SKVGTAVAGTGFEFIWKVLQYDTGLPTTLERLSPKIPSVPISGEDSKMEVIAESGKGVQNI
 IAIAADQLRGSNNIVGGGTRRAIQQQQQQQQETQAVVPVNVPARFEPTFTEIELFLQNK
 FRNVIATIIISMMMLVSNEEMKIIKEVCEHVSHIMVDGLYVALDPRKAIEEILERITAEQ
 NGTITDTGNEGYSRLRYASSGRLFINDEASEEAAAAIGGGGALGTGRRVPVELRSILDKL
 NTIGSTTQQQQQQRQQRQANNNTVPEDIKVHNEQMQKIRDSSLFTSKLLNYIRDDGRK
 DRIKTNISETLKKYSRIPSYFIASKAQKPIPWKHTKDNINLNKIPEDLNFSPAQNLFVPV
 NPRHILDMQWLNCISIIETATRDSAIVMQSFQEQADKTTTQLEELLSQWNNIVSQVTDE
 KSPAYVSSVKLEWLNNEASRIAAIRENSEKSKIVMGVQKIVNIDELGIVAVARSIVDVD
 FYIKMPNVWASRDWKNLIYYAVNIAATPLINNISRIGIMASQTSVLYDSSLALIAAEQAT
 RNITM

(SEQ ID NO: 235)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| ST20_YEAST (Q03497) SERINE/THREONINE-PROTEIN KINASE STE20 (| 46 | 0.001 |
| YN23_YEAST (P53832) HYPOTHETICAL 52.3 KD PROTEIN IN MRPL10- | 41 | 0.047 |
| Q24523 (Q24523) BUNCHED PROTEIN, CLASS 2 ISOFORM (SHORTSIGH | 40 | 0.11 |
| NIT4_NEUCR (P28349) NITROGEN ASSIMILATION TRANSCRIPTION FAC | 39 | 0.24 |
| O62235 (O62235) F36F2.3 PROTEIN | 39 | 0.31 |
| MYSG_CHICK (P10587) MYOSIN HEAVY CHAIN, GIZZARD SMOOTH MUSC | 38 | 0.53 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 234 to 1099 of CT619: this corresponds to nucleotides 165434 to 166299 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=1920

5'stop=1941

3'start=2022

3'stop=2045

5'primer=GCCAGTGAAAGTGGATGAGACC (residues 1920 to 1941 of SEQ ID NO:234)

Tm5=59.57

3'primer=AGTTGTCTAGCCTCCTTTACAGCC (residues 2022 to 2045 of SEQ ID NO:234)

Tm3=58.24

probel=CTTGGATGTACATGCACC (residues 1943 to 1960 of SEQ ID NO:234)

probelstart=1943

probelstop=1960

direction1=Forward

Tm1=69.03

score1=1.96

length=126

WO 01/38351

168/201

PCT/US00/28888

CT1020
 Nucleotide
 Genomic coordinates:
 Start: 62990
 Stop: 63659 (SEQ ID NO: 236)

Amino Acid
 MEWINQRTSREDLFNTYTGNVIRSAAKQALAEKHAERERGEKAWTTSAAAAASSNFNN
 VQODYTDDITQVSIANSVLNNPFLKRYAKLIDNLAISLPPDIEDDVIHTRDASNSTV
 RVDGANIYFAIIDGDLGVYPKQYISDKVLCGSLNREKALFYNSKKNKWTYGCNLNFDIVD
 AAIMKHPDYKEETTSTKHIRKILGIGASEKLNITHYLNFIQ
 (SEQ ID NO: 237)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|--|-----------------|------------|
| P75338 (P75338) MG307 HOMOLOG | 32 | 2.7 |
| Q80990 (Q80990) LATE MINOR CAPSID PROTEIN L2 (FRAGMENT) | 31 | 7.9 |
| Q18668 (Q18668) HYPOTHETICAL PROTEIN C47D12.2 | 31 | 7.9 |
| Q80991 (Q80991) LATE MINOR CAPSID PROTEIN L2 (FRAGMENT) | 31 | 7.9 |
| Q80985 (Q80985) LATE MINOR CAPSID PROTEIN L2 (FRAGMENT) | 31 | 7.9 |
| Q02689 (Q02689) HYPOTHETICAL 45.6 KD PROTEIN IN COI INTRON | 31 | 7.9 |

Comments:

TaqMan Primer/Probe Sets:

5'start=331
 5'stop=354
 3'start=379
 3'stop=402
 5'primer=CACACTAGAGATGCCTCCAACCTCT (residues 331 to 354 of SEQ ID NO:236)
 Tm5=57.74
 3'primer=ACCGTCAATTATGGCGAAATAGAT (residues 379 to 402 of SEQ ID NO:236)
 Tm3=58.84
 probel=CAGTCAGAGTAGATGGAG (residues 356 to 373 of SEQ ID NO:236)
 probelstart=356
 probelstop=373
 direction1=Forward
 Tm1=69.01
 score1=1.98
 length=72

CT1021
Nucleotide
Genomic coordinates:
Start: 69264
Stop: 76212 (SEQ ID NO: 238)

Amino Acid
MSDTGQMEENRPATQKRRPGDEEEETGSSNPYYANFGDDATYSMYTGEGKRGKFVLEP
PKERSVQRVQKPPKEKEEREQRSNVTRRPGQEFQKVLQDRSRERSEKLGQNLAEKGLQ
ERQKKYTPKVAQMTMKKIIRFREGGRKFAPQQQTSKGAATNVLEREEIEMAAEREQPV
EITGDTILGGLGEEDDEDMGEDELTIQHSSMAVSQPVQQIVVSSPIPPKPTRPAPDIPQ
EDIVGKNISQLPPLPLDDYEDDEDEHLYEEVNDFLVAPPTAAAAASTRPPRPNI PPPPP
VVAVADETLKNLASIAALEKEAEEQRAAAVEREREVEEQRAAAAAAAAAAAAAQREADEKR
EREAEEQRAAAAAAAAAAAAAQREADEKKEREVEEQRAAAAAAAAAAAAAQREADEKREVEE
QRAAAAAEREILAQQLQEMKEQMRIKEEERRKELADKEEEKRRELAKEEEKRQEILAKE
EQLEKLNFLQGTETISKRALEQMLEEKASRSRASAQLAIQAIEYEDELPAVEPQGG
LVPMDDTLYGKMYDLNKKLEVQNNTLSAFEDVNKTNEQNQLVAQSLEKSAKAEKLTSSQ
KHLPVDDPAFMQRIITERDFSLKNLGNVYKRVLGVTFTLKRDLFKSKALITDKESRDLEV
RLTDVSTDLRANLNTILERLDVSVNIRSGGTLYTKFTEADTALADQVPSRIEISNRSRS
ALLPFSSAGLDTNFTNSSDKYNEIVNLSSINEAMNILENIVPTLNQIKIDVTNLLTVS
SSROYAIEERVYSDVSRMDSEIRKFLAIMNSKISPYFKGWDTERQRSIADSISSQIKSN
DKIKESVATLHDINTTSRIRSNPLLHKSSVLSSPDLNANDFRNFDIQQGSQFTYDVL
SGQNIDDLSLASKTTEKVTCLCLESILDVHKNALSLNLPATYPAGETSMEESSGLA
VDIRQEIGKNISDSSAELSRITLSEALQIFQQQQQQQQQFQQQQLLQQQQDQNNQQQQLLQQ
QIEEQQRVQEQQQQQQDQQQQEQQQQEQQQQQQEQQQQEQQQQQQEQQQQSDQFRQQL
LQQQQQFQQLLQQQGRRRGGDDGDEEREEREEGAEDDVCVRKVAESVATKYTADLTTLFQ
REENNFQSKIASAKGLTLVFATPPSPIMNLTKLREEYSTFTTQCFSKLTAENNSIMRIFP
ERIVEVCKSKNINLMGKLYLIITTAQTEMEDRVKNILSGIFNQIEEFSNNVKQQQQQQA
SASSTNPPPPSTPSTPPVTSMQVCELDQRTLEKAAIVEAITLANAVLQTTKSASAPST
AAEREIALKLENGKTSIRMEKVDLSSGATGVSDQQKWIDESTSKOELEDFIAEENFVETA
HNEMDIGLILDAKKNDPTRDANLRLVKPHGINVQSFYVLRVTLWGETDILDEDTVHPEY
FRQYIDRNWKVEEHEREDTLKALGVSLSDTLAHIKDYSPSVKNDASKSVFPALNTLLYN
IFAIIDGGMISSLSRTAFIYRKFLRQSMTDKEVAQGPVRSQCEATIASLFTACSNLLRSS
PLADKVEPRLOEKLA AAAA VDTSTGDMFRIRVCHLMYNFIVAVNLCNNRINYTLNVLRA
SGLANKKVAGKTTKGHTSSSHRFGSYDVTYDFSVLYKILQLQKQNISLLEKGFNAWES
CVAAMAAFTADPSLSISDADQSILFPLEGGEIVIEKHENDAENKNDVMQELWKETALTLM
AKELNSYYNWFYISKDTMEKLARVCRMIIIGIVKAVLRLTNKAESLVDTNALSDFIKLPV
IPIDDTKTLAINIVVFTLNNVIKPMVSEFKQMRQKDDGVSSAYFSFQNIQQQKHQQTAS
ILDWACAPGKLTAAHVFIISGYENHIKLLKDDLLWGASMKFPADGRGTVEGWAQQYNN
ESVLEDFDTSIEVNAPASGLLIPDPLSSMFGKNGGSSSSSSSKDNTIIGKGLILNR
QVVGQEQAPPINTSSDTKKIRRDANIEPIIGTPYSVIKASKGVSSISVLDDFNEDSPEDFA
LKTSIINDAIREIGQRMITYTRPIFDHQTQKNIHYSSPKIILEGSDLKNGQRSGQSWAPSS
SSLTLASDWNLPSELLEYRELATKQVEKEEEKSEREEDKGQKLNEKLSFVFNKAIGTIQ
QQHQYSERGGGMKRYQQHSADQASNGGIDDIELMNSKDATSMRKAKLALAVTNKIAAAAA
RDGENSSAKPSNFGNRLDEAINPGALLRRGGGVRRGGQTPQSSMLTMFRPGQTGGNSSWW
TTNTPLIQRRTTSVGNLNLVVLVPLNLLDSHPPTFN
(SEQ ID NO: 239)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|--|-----------------|------------|
| 076853 (076853) SRF RELATED PROTEIN | 136 | 2e-30 |
| AAD50121 (AAD50121) ADENYL CYCLASE | 131 | 4e-29 |
| 077033 (077033) TRFA | 120 | 8e-26 |
| AAD46501 (AAD46501) LATENT NUCLEAR ANTIGEN | 120 | 1e-25 |
| O88542 (O88542) OPA-CONTAINING PROTEIN 1 | 118 | 4e-25 |
| Q62006 (Q62006) OPA REPEAT (FRAGMENT) | 118 | 5e-25 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 462 of CT1021: this corresponds to nucleotides 75184 to 75645 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=3774

5'stop=3799

3'start=3851

3'stop=3872

5'primer=AGCTGCTTCAGCTTCTTCTACTAATC (residues 3774 to 3799 of SEQ ID NO:238)

Tm5=57.20

3'primer=CGTTGATCATCCAACTCACAAA (residues 3851 to 3872 of SEQ ID NO:238)

Tm3=57.20

probel=CTGTTACAAGCATGCAAG (residues 3833 to 3850 of SEQ ID NO:238)

probelstart=3833

probelstop=3850

direction1=Forward

Tm1=69.00

score1=1.99

length=99

WO 01/38351

171/201

PCT/US00/28888

CT1022
 Nucleotide
 Genomic coordinates:
 Start: 79064
 Stop: 83375 (SEQ ID NO: 240)

Amino Acid
 MDFEGTTSSTPSKMSQLYSSVKKVAEHSFANLHDKATLASKVIKDLGERKKMSTPKSSS
 DGQKLDKAMLDDIINEYQAVKSTADNSIESTIKEIENVLESVRRTKIESEAKNSVTSSPE
 KVFSVEDLEIYSKGRVCKGLKLNANCSRIGGKYAVSMSIKKHNVSSFENNNNQVFSEEP
 DCFMLETITYPLVGFTSTEDGNTYAVFLTGVGLERSLPKYVPVFDMNAGIQTLMNTGLRM
 AKLPVLKMFGRTEYDNLEDFYITSIETQSFDEEENDARMRCHTEDLERKKRMNDAPAITP
 HVAVYDYSGDGKEQLLYMITEYENTASWCNANGVVTSDSGFSNECAISDMNDLCCFADCI
 DVTVNNEEHEERSMNIVVESDRRLFDASPSPIKTEEDGENSSSSSSSPTVPPTPYEGNA
 VVEGEEEEEEIDEDESSKYEGSEDALVMKKLAKLSTMKQMRVKNPALKITSGGNNSSS
 SINNEDDGDDDDAVDATALCPQTEATVKNSFMAPNDERTENILYETMQISLAKICNNPSS
 MSSYRVFTNKLQECINTMDDSI RRRPTIWTEESQQFAKGLLFDEVVTSIVAHQMAQDICK
 SEIFGGMFNANSTNIKGYEGQKKSLYGNKHISSSCFKTNTESNVNNALFAWVSKSLHSG
 TVIPNVFSFKMASEKPSKMKRKRRTSSASSSNDEHQEPSTKMMKNDEGEKVQESSSPSSS
 STPEQQQAGHDKETINLIPLSFIKMPRSNVNGSASYLSEIFGQRLCGLSDASSTFKRMC
 KTFEDLENEIMRSSFTRLTRYEREVTRLYEKCRSQAVDIEENEMDVLSHQGELEFAEFLED
 PIAYFEEVLENIKSWSLENVNTPKRKNKYAKVLVSVNAIRRTYEEYHAFSKFVPMFLFNL
 IKRELEGDNYTHDVHFSSTCLWYLTVMTRNRICDVLQYINNNNDNEETDIVEEEEEEGEG
 EEDKMEESMDVEQQKQVRKGGKKGQKFNSIGDQVIRKFVKSLENSMVVSIAINSLISG
 ISWMNKKIPPGFLKDSSTINTLDEVSRFVFSVDVKINRKINGTDDKYETVFGVSTRVDSHI
 VGPFPSIPVDFSSAGLDKASCGLYVNTIDGKGILTISPKYDSLNDDEDVDSTTDKLEKDI
 LHLSKHDTFNINKNKVLPFYNISPSSSLTEKKKTKFNRKKISSGMSNNNGMCVQTPSSS
 NSVSSVSIVAPSSSVLALSCSLSTKKSIIWNNMFLTSRNMWRCGFVVPKLCSEFIVN
 HRHAVKLVAETAPKTKLCRNIIIDNRKIRFNGLLKKVCKSVSAFTGESTYLLNKNMTATSP
 SDLNLCIYTSSLNDPLYTCKLTHEEYQDGNALDDYGAVFVNYTFKSIKSCSSKDETTADDN
 AAAADDDGTTSTSSSTDTDAAAIQDFMHVMIKKIDAMKDIRGKYKKS LAKKTKKH
 (SEQ ID NO: 241)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| UBP1_YEAST (P25037) UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 1 | 50 | 1e-04 |
| O97236 (O97236) PFC0230C PROTEIN | 48 | 3e-04 |
| CAB43859 (CAB43859) HYPOTHETICAL 85.2 KD PROTEIN | 46 | 0.001 |
| NSR1_YEAST (P27476) NUCLEAR LOCALIZATION SEQUENCE BINDING P | 44 | 0.006 |
| Q21000 (Q21000) SIMILARITY TO C. ELEGANS MYOSIN HEAVY CHAIN | 44 | 0.006 |
| Q18918 (Q18918) CODED FOR BY C. ELEGANS CDNA CM11B12 | 44 | 0.006 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 1010 of CT1022: this corresponds to nucleotides 82212 to 83222 of the genomic reference sequence.

WO 01/38351

172/201

PCT/US00/28888

CT1023
 Nucleotide
 Genomic coordinates:
 Start: 93228
 Stop: 94137 (SEQ ID NO: 242)

Amino Acid
 MVSSITHLSLLFVVAVVASVVFTTEGASVRVKRCVSPCPDVIDPDHRCQGRLCRRSTRG
 GDDDDDDDDGGTFTDVGSGILGRKKRAAPPEDEEDDFYRKKRAAPPEDEEDDFYRK
 KRAAPPEDEEDEFYRKKRAAPPEDEEDEFYRKKRAAPPEDEEDEFYRKKRAAPP
 PEDEEDEFYRKKRAAPPEDEEDEFYRKKRAAPPEDEEDEFYRKKRAAPPEDEE
 DDFYRKKRAAPPEDEEDDFYRKKRAAPPEDEEDDFYRKKRAAPPEDEEDDFYRK
 KR
 (SEQ ID NO: 243)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q18401 (Q18401) COSMID C33G8 | 113 | 1e-24 |
| Q9YTL7 (Q9YTL7) ORF 48 | 98 | 6e-20 |
| P73032 (P73032) HYPOTHETICAL 185.1 KD PROTEIN | 95 | 7e-19 |
| Q43687 (Q43687) EXTENSIN-LIKE PROTEIN (FRAGMENT) | 88 | 7e-17 |
| Q09085 (Q09085) EXTENSIN CLASS II PRECURSOR (CELL WALL HYDR | 83 | 2e-15 |
| O14686 (O14686) ALR | 77 | 1e-13 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 214 to 1051 of CT1023: this corresponds to nucleotides 93435 to 94272 of the genomic reference sequence.

WO 01/38351

173/201

PCT/US00/28888

CT1024
 Nucleotide
 Genomic coordinates:
 Start: 94623
 Stop: 95742 (SEQ ID NO: 244)

Amino Acid
 MDNLKGEFVALKTDLTHYKTQLDRSILVFVDVVGRLYVIVNSEQTAKKEGLATRVAKQAT
 EIQQFKDEINNKNALNTLDDIIYIFDHGGSFKRAKHKAIEAREYSKPLRELECMFTR
 IADMLTLTFMTVYTNIIITEFRHSSEQATNSINVTLGRLFLCDDLQNQLPKEEEEEEDLKQ
 KFITFHANLYMLDTRLKKDLIIFKDVIQQLHVILQKDTYAVKEGVAIRCAQMNEISQYR
 DNLKDNYNFTFSNILEIVYIFDHGGHFEEVKHKAITLTRNYLKTLMGLKCMFKRISEMLS
 LTFLT VYTNVIAEFINASNISDREINNYLVQLVTCNELCNQLPKPKQYRPLSLIDNIAFY
 SLSVQKHLSGFL
 (SEQ ID NO: 245)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| AAF04635 (AAF04635) HYPOTHETICAL 43.2 KD PROTEIN | 747 | 0.0 |
| O96219 (O96219) HYPOTHETICAL 139.4 KD PROTEIN | 39 | 0.051 |
| O64554 (O64554) YUP8H12R.45 PROTEIN | 39 | 0.067 |
| TPR_HUMAN (P12270) NUCLEOPROTEIN TPR | 37 | 0.20 |
| Q99968 (Q99968) NUCLEAR PORE COMPLEX-ASSOCIATED PROTEIN TPR | 37 | 0.20 |
| O97291 (O97291) PFC0960C PROTEIN | 37 | 0.26 |

Comments:

Hit to public SBV sequences:
 gi|6165655|gb|AF099142.1: CT nucleotides 1 to 1119 match nucleotides 3220 to
 2102 of the public sequence with a 100% homology, a score of 2218 and an Evalue
 of 0.0

TaqMan Primer/Probe Sets:

5'start=478
 5'stop=499
 3'start=576
 3'stop=598
 5'primer=TTGTGTGACGACTTGTGCAATC (residues 478 to 499 of SEQ ID NO:244)
 Tm5=58.28
 3'primer=AATCTTTCTTTAGGCGTGTGTCC (residues 576 to 598 of SEQ ID NO:244)
 Tm3=57.55
 probel=TGCGAACCTATACATGCT (residues 558 to 575 of SEQ ID NO:244)
 probelstart=558
 probelstop=575
 direction1=Forward
 Tm1=69.00
 score1=1.99
 length=121

CT1025
Nucleotide
Genomic coordinates:
Start: 95824
Stop: 97369 (SEQ ID NO: 246)

Amino Acid
MDSCCLISRITPELAGKLTWIFIPENNFKIVQNSLPDDQVISQFRYFDHRHCYTFMEILM
ANIKIQDRKQNTTAICELTTGREGLLCRRTIPVFLGSEEKREELLGNLPEGAEIFRPREV
MQVIGTLLDKKLEIDDGIASVKAALCAGSSSLYLIMSHIVKMTFSAITNMKDINEEYFVD
FIFRHKQFLNPEFFKHLISLLKNSRKEHVAHLVRRLEHFLMLWTLSKMRFTEMEENYFPI
SSDSYGICEKCAKRTPKYKLRIFRERKCCDRCCRLYHQPPPEVYNWDGKITQQSNKGY
INAGDEIIIGMLNSNDKGKTFPPIPKMVVRRVVDGVYGGGTILSKILKFRQANIPTCLFVT
CNKCNRIFRILTILGPTRNILCPPCRKKSVAVNTQQKGENKPSFVQKGTKRLRVDTGSNKN
TLEKFCSWERFNTFVLLPWLGYTIESKWQNWESFLGYSSSTRYKELWAFVKNQEISSMKDS
YIKIEDIDQLLSILQDQKGVFETVCKIKSRDGL
(SEQ ID NO: 247)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| AAF04634 (AAF04634) HYPOTHETICAL 45.2 KD PROTEIN | 689 | 0.0 |
| STAS_MOUSE (P42230) SIGNAL TRANSDUCER AND ACTIVATOR OF TRAN | 32 | 7.5 |
| AAD20715 (AAD20715) PUTATIVE DNAJ-LIKE PROTEIN | 32 | 7.5 |
| CAB60246 (CAB60246) HYPOTHETICAL 80.9 KD PROTEIN (FRAGMENT) | 32 | 9.9 |

Comments:

Hit to public SBV sequence:

gi|6165655|gb|AF099142.1: CT nucleotides 1 to 1545 match nucleotides 2019 to 477 of the public sequence with a 99% homology, a score of 3031 and an Eval of 0.0

TaqMan Primer/Probe Sets:

5'start=672
5'stop=693
3'start=765
3'stop=786
5'primer=CCTTTCCAAGATGAGGTTTACA (residues 672 to 693 of SEQ ID NO:246)
Tm5=57.77
3'primer=ACGGAGCTTGTATTTGGGAGTT (residues 765 to 786 of SEQ ID NO:246)
Tm3=57.84
probel=TCCAGCGATAGTGATTAC(residues 721 to 738 of SEQ ID NO:246)
probelstart=721
probelstop=738
direction1=Forward
Tm1=69.03
score1=1.96
length=115

WO 01/38351

175/201

PCT/US00/28888

CT1026
 Nucleotide
 Genomic coordinates:
 Start: 97547
 Stop: 98789 (SEQ ID NO: 248)

Amino Acid
 MESIKLFTVAGLNMEQANQVAEEIKSEYKTEEEKRIAQEVFDKFTKKLIMQVDTSKHLLT
 RENPNRFVSRPIVHEDLWEMYKKEVACFWTLEEIDFERDPKDWEKLTQDEKDFILQILAF
 FASSDGIVIENTLTRLRQVAQIPEARSFFDFQVGMESIHGNVYGELIDRLVPDEKDKAIL
 FNAAQHFPFAIKKKEQWAINWMSNNDLAELIVAFAAVEGIFFGAFASIFWIKNRGILPG
 LTSSNEFISRDEGLHRDFACMLLKKGFVDTPSRERILEIVTEAVRIEQEFLTVSLPVKLV
 GMNCKLMSQYIEFVADKLLVEMGLEKHYNVTNPPFFMDNISLENKTNFFEKRVAEYQRAQ
 VMASINKIKKDQQTQETGSPLPILTAPPPVSSSSSEQEDVEDGVGDYISYDDF
 (SEQ ID NO: 249)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q9XYN8 (Q9XYN8) RIBONUCLEOTIDE REDUCTASE R2 SUBUNIT | 388 | e-107 |
| RIR2_HUMAN (P31350) RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE M2 | 381 | e-105 |
| RIR2_MESAU (Q60561) RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE M2 | 381 | e-105 |
| RIR2_MOUSE (P11157) RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE M2 | 378 | e-104 |
| Q27124 (Q27124) RIBONUCLEOTIDE REDUCTASE SMALL SUBUNIT | 377 | e-104 |
| RIR2_BRARE (P79733) RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE M2 | 376 | e-103 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has sequence identity to nucleotides 1 to 959 of CT1026: this corresponds to nucleotides 97786 to 98744 of the genomic reference sequence.

Hits to public SBV sequences:

gi|7672988|gb|AF144620.1: CT nucleotides 1 to 1242 match nucleotides 1 to 1242 of the public sequence with a 99% homology, a score of 2454 and an Evalue of 0.0;
 gi|6165655|gb|AF099142.1: CT nucleotides 1 to 299 match nucleotides 299 to 1 of the public sequence with a 99% homology, a score of 585 and an Evalue of 1e-169

TaqMan Primer/Probe Sets:

5'start=642
 5'stop=663
 3'start=706
 3'stop=727
 5'primer=TGCTGCAGTTGAAGGAATCTTC (residues 642 to 663 of SEQ ID NO:248)
 Tm5=58.19
 3'primer=AGGTGAGACCAGGCAAAATACC (residues 706 to 727 of SEQ ID NO:248)
 Tm3=58.13
 probel=TTAGTGGTGCATTCGCAT (residues 665 to 682 of SEQ ID NO:248)
 probelstart=665
 probelstop=682
 direction1=Forward
 Tm1=68.98
 score1=1.98
 length=86

WO 01/38351

176/201

PCT/US00/28888

CT1027

Nucleotide

Genomic coordinates:

Start: 98874

Stop: 99789 (SEQ ID NO: 250)

Amino Acid

MNLLPIFLTTFVAVDACSCSTICLLPDGKKQPLVFDVLEEVVYPTDVC GPKGAGELFT
 GVDLLTLCIGGKNNNGEWSGKGPCPRINNAVVERDYSLEEDCKGFRKGFRI PGTDHFHT
 VFSLCWVDRDMHAKWVRNKNPGIVTDDDLVDSGIRTKFKYSSKIFGKGFNRP LYSLD
 YQERIKILKSHFNKRTGNFFARGHLAPAGDFFLASERWATFALEN AVPQIQNHNNGEWKD
 IENRARTTPGAAWAETGPIFYQHKKKEYLDKKKKYIPIPHALYKIVYDKNNKELFRVQSD
 MSWK

(SEQ ID NO: 251)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| CAB55635 (CAB55635) DEOXYRIBONUCLEASE I PRECURSOR (EC 3.1.2 | 59 | 6e-08 |
| Q9Y2C4 (Q9Y2C4) ENGL-A | 48 | 7e-05 |
| NUCG_MOUSE (O08600) ENDONUCLEASE G PRECURSOR (EC 3.1.30.-) | 45 | 0.001 |
| NUCG_BOVIN (P38447) ENDONUCLEASE G PRECURSOR (EC 3.1.30.-) | 45 | 0.001 |
| O73911 (O73911) K123 PROTEIN PRECURSOR | 43 | 0.004 |
| NUCG_HUMAN (Q14249) ENDONUCLEASE G PRECURSOR (EC 3.1.30.-) | 42 | 0.005 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 808 of CT1027: this
 corresponds to nucleotides 98914 to 99720 of the genomic reference sequence.

WO 01/38351

177/201

PCT/US00/28888

CT1028
Nucleotide
Genomic coordinates:
Start: 103891
Stop: 104680 (SEQ ID NO: 252)

Amino Acid
MSLAVTEDYGHNEKLIKRLQTSVYHTPLLGAHVMSISDYIISRRFMNYTNLLKQVEYV
FDEETGAVIANICLLKILERCAQKGGIYDAPEDVAFFNSKMGEVTRLFTIIGGRPNMTVR
VNFKHGQTNNPAYGYLTDDNDTTVTTPPVTPPPSPAARRSPFFTRLISESSSVVDHYVLM
HDNPKRSSFKVYDIHAETFPKAPSVPTFPKTSFEISDVTLDCSMEIFSRDRDVLDNVH
DYIANDVPVFLVDVVHRGSSLR
(SEQ ID NO: 253)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|--|-----------------|------------|
| PAXI CHICK (P49024) PAXILLIN | 36 | 0.39 |
| Q24459 (Q24459) POLYCOMBLIKE NUCLEAR PROTEIN | 35 | 0.67 |
| LA17 YEAST (Q12446) PROLINE-RICH PROTEIN LAS17 | 34 | 1.5 |
| Q14676 (Q14676) KIAA0170 PROTEIN | 33 | 2.6 |
| Q94399 (Q94399) ZK265.2 PROTEIN | 32 | 3.4 |
| O88699 (O88699) HOMEODOMAIN PROTEIN | 32 | 3.4 |

Comments:

TaqMan Primer/Probe Sets:

5'start=390
5'stop=413
3'start=457
3'stop=477
5'primer=TCCTGCCTATGGTTATCTCACAGA (residues 390 to 413 of SEQ ID NO:252)
Tm5=58.77
3'primer=TCTTCTGCAGCTGGAGATGG (residues 457 to 477 of SEQ ID NO:252)
Tm3=59.15
probel=CCTCCTGTTACTCCTCCT (residues 439 to 456 of SEQ ID NO:252)
probelstart=439
probelstop=456
direction1=Forward
Tm1=68.94
score1=1.94
length=88

WO 01/38351

178/201

PCT/US00/28888

CT1029
Nucleotide
Genomic coordinates:
Start: 104759
Stop: 107330 (SEQ ID NO: 254)

Amino Acid
MASTSSSTKKRVHEEDENLIPQPKKKKSKKVLPPFPVDKYRAVDKKVVNLIHKILDQEKDH
LSSTELQMITECNGAREDLLKHLDEGEFNPTIIIEVVSSMPIETIYEILSSSADKKFVQ
ISLSMLIHILFFADKGTMWVSNACVQNVLGNDYKVEFENIRKKYLILEDLLNGVSNHWSE
HGPLSHMLHSSIPIVQDMLLNRLVRYFSTYDGAQFDISFIINSVLWGIDKSVLNELTQL
ISRGVFIVSYVPMRVRTPSKDSNRPQNTPSQNSALGMKLNTFSSRISVYRNNTFKKLTE
LVHNFYDGSKDASSSSPPPPSLSDSVNTFVRLYTNYDIFLKVISDWKMPYGFKKTFDVL
YSKGLMTLSVSEYTLKKELVTFLRALKEREILYKMEKRDIICILKKSFLGFNFRCLKQL
LPFLKHFLLKIEEVKHIAFVFRDYSLMCKTQKDLQSFPAIQSASLFMEFPWLAKTWIDD
DDEGGKGHTLLTFAIVHRYPLISQLISHPILKSLVNTTCRDKHFTPLMHLANTSIMYQC
NTLLCLIINGAKPEFINKFNENVLHIAIENVNYGVITELRGTLSSSEQIEKMVNVRMMDN
TTPLMIALARENIVLAQLFDGLYKPKIKVRFSGSSKRLRIPEFVLLKGLKESVAYLETRNI
SYDINI IKDAVMDNSLFEEYEIAAAGLRGNCDPEADEKTMNTWNFFTKNSTKWASSIF
QKNRQKFVKIVDGMNRTYEDSECAICLDSLGDLP SGRTTCGHCFHNVCWLSLIRMSGPN
NGSRARGGGIKCPCSCQVTC LGKRLGVADYDIETEEERDTKNVVPSEEGRREWRKIGVD
RYEFLVGGVWTNEIKL
(SEQ ID NO: 255)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| Q9Y577 (Q9Y577) RING FINGER PROTEIN TERF | 43 | 0.007 |
| AAD27914 (AAD27914) T23015.13 PROTEIN | 40 | 0.058 |
| AAD56722 (AAD56722) AUTOCRINE MOTILITY FACTOR RECEPTOR | 40 | 0.076 |
| Q9WV59 (Q9WV59) RING FINGER PROTEIN TERF | 39 | 0.13 |
| O82239 (O82239) F17A22.9 PROTEIN | 39 | 0.13 |
| RO52_MOUSE (Q62191) 52 KD RO PROTEIN (SJOGREN SYNDROME TYPE | 38 | 0.22 |

Comments:

TaqMan Primer/Probe Sets:

5'start=1406
5'stop=1425
3'start=1491
3'stop=1514
5'primer=AATTCCTTGGCTTGCAAAA (residues 1406 to 1425 of SEQ ID NO:254)
Tm5=58.29
3'primer=TGGCTTATTAAGGATATCTGTGC (residues 1491 to 1514 of SEQ ID NO:254)
Tm3=57.50
probel=TTGGATCGACGACGATGA (residues 1428 to 1445 of SEQ ID NO:254)
probelstart=1428
probelstop=1445
direction1=Forward
Tm1=68.95
score1=1.95
length=109

CT620
Nucleotide
Genomic coordinates:
Start: 172439
Stop: 171509 (SEQ ID NO: 256)

Amino Acid
MDNLITNDNIILVTFSLGLAVGCSMTIGLALAMNMLVKCIDRTTTCISCSPWEKNKNKKN
RNGSNTLESSFISHVRFNTPDKDLDISEPLKSTTYDLANVTPQVTKLVTFSGPTYASPPT
PRPVANTPQQQPTSTNKEEESVYMPMSSCSSSFSSDNSLPLTPPPSPPRSNGGDYVSYV
NGRHLKLPSPNPPSPIFNIKNEEGEDDNVEEHVYEVPEVPQQSPSIQKCIQELKEMKHHK
NTLTRSSSNNNNNAPRITQVTFKKFPPNNNNMWENHVGNTTIVSSTPSPTFIPSPKSI
RKLSFRRKQ
(SEQ ID NO: 257)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q24035 (Q24035) ENA POLYPEPTIDE | 46 | 4e-04 |
| Q41805 (Q41805) EXTENSIN-LIKE PROTEIN PRECURSOR | 43 | 0.002 |
| Q40786 (Q40786) ARABINOGALACTAN-PROTEIN PRECURSOR | 43 | 0.003 |
| Q9XZU9 (Q9XZU9) LARGEST SUBUNIT OF THE RNA POLYMERASE II CO | 42 | 0.005 |
| O36428 (O36428) GLYCOPROTEIN PRECURSOR | 42 | 0.005 |
| YHC4_YEAST (P38741) HYPOTHETICAL 80.1 KD PROTEIN IN SNF6-SP | 42 | 0.005 |

Comments:

EST confirmation of the predicted transcript:
An isolated EST has equence identity to nucleotides 665 to 1 of CT620: this
corresponds to nucleotides 171494 to 172158 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=494
5'stop=513
3'start=558
3'stop=578
5'primer=CGCCATCTCCACCTAGAAGC (residues 494 to 513 of SEQ ID NO:256)
Tm5=58.49
3'primer=GAAGGTGGGTTTGAAGGAAGC (residues 558 to 578 of SEQ ID NO:256)
Tm3=58.68
probel=GGCGGTGATTACGTGTCA (residues 517 to 534 of SEQ ID NO:256)
probelstart=517
probelstop=534
direction1=Reverse
Tm1=69.00
score1=1.99
length=85

WO 01/38351

180/201

PCT/US00/28888

CT621
Nucleotide
Genomic coordinates:
Start: 173054
Stop: 172505 (SEQ ID NO: 258)

Amino Acid
MGESIFDAVSLATNNPKKSNSRNKKLLRELKNMRKDFPSTFLQCRMIDFHFSGDIIDKHY
CHSVNVPDVPNTIFAVFLPEEDRANNPGLYDSIEGVCITVEQGELCIINKSSVHEFNIL
VSLHKDLFGEDILDGIETASREESRSIHLYLEAGQSIRTPIPRPEGTNTVNYTIVFSNQV
TV
(SEQ ID NO: 259)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|--|-----------------|------------|
| O88632 (O88632) SEMAPHORIN IV ISOFORM B | 34 | 0.54 |
| O88633 (O88633) SEMAPHORIN IV ISOFORM A | 34 | 0.54 |
| O29244 (O29244) ABC TRANSPORTER, ATP-BINDING PROTEIN | 34 | 0.71 |
| O54948 (O54948) SEMAPHORIN IV HOMOLOG (FRAGMENT) | 33 | 1.6 |
| Q13275 (Q13275) SEMAPHORIN IV PRECURSOR (SEMAPHORIN III/F) | 32 | 2.7 |
| INA9_MOUSE (P09235) INTERFERON ALPHA-9 PRECURSOR | 32 | 3.6 |

Comments:

WO 01/38351

181/201

PCT/US00/28888

CT622
Nucleotide
Genomic coordinates:
Start: 183817
Stop: 180274 (SEQ ID NO: 260)

Amino Acid

MEAASNLRITEGAGVLDIDNEDDINNNDVDSNLYEDEDEEEMNEDEEEEEEDYEDDE
DTGVRNNGRNKDPSSKKQSKFVRDVTNDMYDDDDDEEEEEEEEEDEEGEGGEYDGNLED
EEEEGDEYEDDNEGEDEDEADPALLALAAQEDATIIPENQWKSIVNTPSPVGPNRQVL
PMLNLFLENVAMGGGSAEEQKNKEDDNQIEPVEEEEEEEEEEEEEEEEEEEEEEEE
EKEPIEQEKNEPEKDEDAIENESVHSHRVESPMSEGGNDGMDYFFSSIAGGNDNEED
EEEEEEEEEEEEEPQKSEEHVETKESVQSHTEYIEEEEEEEYEEYEDSRHTLEDEEIS
TMHQFNNAVRVRSPPPDIECEDAVVFPPIMKETDILPQIKEPSPKAPRMFSLGSGGE
EQYDQLNDIAPPVPSIVTFPPDNEMGEESRDIMDQDSMLMPPPPPPPPPPHQPPLKLP
TNILLPPPPPPPTNQSLSFNNNNNPSFLSTVVGVKNNTLGKKEAERLHKTMSIILKTR
VKTLLETTKNLQCSELVKVVFQDPENPVKPKSEKVMERLKNIIAAELTMKAFLDSAAVTDI
KSAELFRKTNEKLELFQRKQIMSNPLFSAAYASTYIMGERASKIRPSTPAPSLKKVESIS
ELNEDETSMSSSAGGVCAEGDESIAGGGGGGGGGGGEVVEHSSFYSNQTQANLHMEINI
LKEDDDNQPCQTYKLGQRLAFLNNLISFKTSSAVSWSRLVNLSDIVTKASVALFGDTNK
AQEDFEKHQETNDVSDLTSSKQMSKESANIMEEMGLSIGAICFGAISTIEKHI
NKLCDVGRLTIFLNIPIVLLNWPKEFTLSKDYKVLNLDSSSSSKMAVPPPIYVLNSIQ
FDKAVDEEDEDGNGSEAEKRSEDGNMFSEKDKKEAIRRVYDNIRYGDSDNDRSLNHFFGD
AYSGVSNNSKNSMFDLQTQGGGRFGVAYSAGSSIIEHRSPIFDNALNTLVNFMCKRHL
LSAVVIKLLKKAKLSIEVYCIKYKLNQASEKYNKKKGKHKSTSVVPMRNLMYRPSKNQDV
SPSTPAAATAMDVPSVSSHVGRKRTFSFSNDINSNMSSASSVYIDQESSTPSRRRTFMD
LLNNKSSVNSLAKQVKRMKHTKYNSSSSSEDDDEDDQYE
(SEQ ID NO: 261)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|--|-----------------|------------|
| Q9YTL7 (Q9YTL7) ORF 48 | 154 | 3e-36 |
| O40947 (O40947) ORF 73 | 148 | 1e-34 |
| Q98148 (Q98148) ORF73 HOMOLOG | 138 | 3e-31 |
| AAD46501 (AAD46501) LATENT NUCLEAR ANTIGEN | 136 | 1e-30 |
| VG48_HSVSA (Q01033) HYPOTHETICAL GENE 48 PROTEIN | 133 | 6e-30 |
| Q91255 (Q91255) NF-180 | 130 | 5e-29 |

Comments:

TaqMan Primer/Probe Sets:

5'start=1940
5'stop=1962
3'start=2039
3'stop=2058
5'primer=CGACTCCTGCTCCTTCTCTTAAA (residues 1940 to 1962 of SEQ ID NO:260)
Tm5=57.98
3'primer=TCCAGCAATAGACTCGTCGC (residues 2039 to 2058 of SEQ ID NO:260)
Tm3=57.66
probel=GTGGAGTATGTGCTGAAG (residues 2021 to 2038 of SEQ ID NO:260)
probelstart=2021
probelstop=2038
direction1=Reverse
Tm1=69.00
score1=1.99
length=119

WO 01/38351

182/201

PCT/US00/28888

CT623
 Nucleotide
 Genomic coordinates:
 Start: 196292
 Stop: 195506 (SEQ ID NO: 262)

Amino Acid
 MSNGATISDERLILILDKIVERRGVSNLSELLIHPITKHINELLKNTVRHGDVYMKDAE
 LDVRSRLEDIKKDCVLKAIEKQGIDVRQIITDYLAKRKLTQNLVHWYRPPISCTDIDEKI
 QQETGQVGRCSVATYNLRIGGDDGEFTRYDFSIPLGDFKITAKLFRSINDEDVDAVILVS
 RSDVVNDVLSFEAFNRTGERVVIFNFVIVEGKSKDIDIVCKSRYPKHITHILNGESATYAVK
 RIKRGDTRDDILFAITAFKEE
 (SEQ ID NO: 263)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| PYC1 YEAST (P11154) PYRUVATE CARBOXYLASE 1 (EC 6.4.1.1) (PY | 33 | 1.9 |
| BAA85008 (BAA85008) ORF3P | 33 | 2.5 |
| BAA85073 (BAA85073) ORF4S | 33 | 2.5 |
| PYC2 YEAST (P32327) PYRUVATE CARBOXYLASE 2 (EC 6.4.1.1) (PY | 32 | 5.7 |
| O27375 (O27375) ADENINE PHOSPHORIBOSYLTRANSFERASE | 32 | 5.7 |
| Q88444 (Q88444) GLYCOPROTEIN PRECURSOR | 31 | 7.5 |

Comments:
 5'start=282
 5'stop=304
 3'start=383
 3'stop=402
 5'primer=GGCTAAACGAAACTAACGCAA (residues 282 to 304 of SEQ ID NO:262)
 Tm5=58.58
 3'primer=CGTAGCAACACTACACGCC (residues 383 to 402 of SEQ ID NO:262)
 Tm3=57.92
 probel=CCAATATCTTGACAGAT (residues 328 to 345 of SEQ ID NO:262)
 probelstart=328
 probelstop=345
 direction1=Reverse
 Tm1=68.97
 score1=1.88
 length=121

WO 01/38351

183/201

PCT/US00/28888

CT624

Nucleotide

Genomic coordinates:

Start: 195503

Stop: 194651 (SEQ ID NO: 264)

Amino Acid

MSSNGDEPAVTEAEIASVEAQLGAAHHDNSWITRKSDQLKYRLGAIAYSVAKNASIKYIE
 DQVRQEINSHLTNVMTFEHLIEDAFNPVCEAIFKGI PVVMEKVYDVNRRIMEPREDFI
 TEILKEERWRRYIPGFYHTSFSFKYNTIAFTDSSTSFSPINDKHMLSITPPGAAQGDLI
 DLSLSFKIDSSAKTLTLEFNKRSTFAGIVNRPKSVVILSNLRNSDSSDNIGDYLKRNDPI
 YISHDTNGIINPSEDSASLTIHMPEIENASDDLYIDFNLFVF
 (SEQ ID NO: 265)

Top Blast Hits

Sequences producing significant alignments:

| Score (bits) | E Value |
|-----------------|------------|
| 36 | 0.42 |
| 33 | 2.8 |
| 32 | 4.7 |
| 32 | 4.7 |
| 32 | 6.2 |
| 32 | 6.2 |

Y687_METJA (Q58100) HYPOTHETICAL PROTEIN MJ0687
 Q9Y717 (Q9Y717) CAGCR3 PROTEIN
 Q83457 (Q83457) FIBRE
 CAB53329 (CAB53329) PUTATIVE FORMYLTRANSFERASE
 PGK_LACDE (O32756) PHOSPHOGLYCERATE KINASE (EC 2.7.2.3)
 Q9XEQ1 (Q9XEQ1) TNP2-LIKE PROTEIN

Comments:

WO 01/38351

184/201

PCT/US00/28888

CT625
Nucleotide
Genomic coordinates:
Start: 194629
Stop: 193327 (SEQ ID NO: 266)

Amino Acid
MALSNNGGIYIVFAVIVLVIGASIALFFAISGVGKGLHSNAKTKKSKKYKLDISKYTDDD
EKTDNNDNNNGGGGGTVDVINETALQRQTRHFARTLEKAEDEFFTCLADQEFDTYKSE
NVWLIKDKITDGKVSIPEDINVPDVGQAIADENLFDLIGNHDEVKETMDEVVAQKSTN
ITYEQLVIDLTNILLFGTVTVDPDSDENGDESQRSTDPDAEMVMLTTTPSSQLARQQPP
QTPDYLRYSKELVINNIRGGFISDRDMRTWQGRMSVHVNMQRTFNVISAAATNLDLSQ
VGLPEVLQKQGRAAVGGRIEKARIEFSFVVEGNRVRVYATNKTEDCFCSLLPNCYNVKA
SDYWISSASTAKEKTYLFIANKNDETSFFYNFEEGVVEIDLDFMTIDCAPNLPFIKNLP
RPITDNNIMVALS
(SEQ ID NO: 267)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| O96148 (O96148) HYPOTHETICAL 57.8 KD PROTEIN | 38 | 0.14 |
| TRI1_ECOLI (P14565) TRAI PROTEIN (DNA HELICASE I) (EC 3.6.1 | 37 | 0.31 |
| Q43423 (Q43423) CYSTEINE PROTEINASE (FRAGMENT) | 36 | 0.40 |
| O15696 (O15696) ORNITHINE DECARBOXYLASE | 36 | 0.69 |
| MIX2_XENLA (Q91685) HOMEBOX PROTEIN MIX.2 | 36 | 0.69 |
| BAA84604 (BAA84604) MYOM PROTEIN | 36 | 0.69 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 351 to 1 of CT625: this corresponds to nucleotides 193270 to 193620 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=651
5'stop=669
3'start=729
3'stop=748
5'primer=CCCAGACGCAGAAATGGTG (residues 651 to 669 of SEQ ID NO:266)
Tm5=58.66
3'primer=ACCGGGCAAGGTAATCAGGT (residues 729 to 748 of SEQ ID NO:266)
Tm3=59.29
probel=ACAACAACAACCTCCTCA (residues 705 to 722 of SEQ ID NO:266)
probelstart=705
probelstop=722
direction1=Forward
Tm1=69.01
score1=1.98
length=98

WO 01/38351

185/201

PCT/US00/28888

CT626

Nucleotide

Genomic coordinates:

Start: 228196

Stop: 227989 (SEQ ID NO: 268)

Amino Acid

MSDMTRNIIIVGLAVVVIALSMVAFMLSVTPALTGFLGLGVLSALGVTLFGCPTMKSPGGG

NATINPVA

(SEQ ID NO: 269)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| SUGE_PROVU (P20928) SUGE PROTEIN HOMOLOG | 32 | 0.35 |
| Q48546 (Q48546) INTEGRAL OUTER MEMBRANE PROTEIN | 32 | 0.35 |
| O86349 (O86349) HYPOTHETICAL 27.4 KD PROTEIN | 31 | 1.0 |
| Q9Y9X7 (Q9Y9X7) 604AA LONG HYPOTHETICAL CARBON STARVATION P | 30 | 1.8 |
| O69867 (O69867) PUTATIVE TRANSMEMBRANE TRANSPORT PROTEIN | 29 | 4.0 |
| O26370 (O26370) HYPOTHETICAL 5.9 KD PROTEIN | 29 | 4.0 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1032 to 768 of CT626: this corresponds to nucleotides 227951 to 228215 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=60

5'stop=80

3'start=151

3'stop=173

5'primer=CATGGTCGCTTTCATGCTTTC (residues 60 to 80 of SEQ ID NO:268)

Tm5=59.30

3'primer=CCTGGAGATTTCATAGTGGGACA (residues 151 to 173 of SEQ ID NO:268)

Tm3=58.94

probel=TGTTACTCCTGCACTTAC (residues 81 to 98 of SEQ ID NO:268)

probelstart=81

probelstop=98

direction1=Forward

Tm1=68.95

score1=1.95

length=114

WO 01/38351

186/201

PCT/US00/28888

CT627

Nucleotide

Genomic coordinates:

Start: 234330

Stop: 233778 (SEQ ID NO: 270)

Amino Acid

MFQKWFESFLDSSRPRLDTTCVCSVYSYFSPCRKHIKFSTSHSHEGIKIHPPSILNHNT

SSPTSGKMCNHHHKRLYLSTDDHTRWYDKNTSCIYLEDIGGVQFMVYEFHLTPKNNQLFS

FPVHLQIHNRNTEKTSLLVFENEEDMRVRNIHPKSKILIPVSKDTVLVENGFRYKVKIVL

SNK

(SEQ ID NO: 271)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| TRBH_ECOLI (P19381) TRBH PROTEIN | 31 | 6.4 |
| BCL6_HUMAN (P41182) B-CELL LYMPHOMA 6 PROTEIN (BCL-6) (ZINC | 31 | 6.4 |

Comments:

TaqMan Primer/Probe Sets:

5'start=187

5'stop=207

3'start=233

3'stop=253

5'primer=CCCACCAGTGGAAAGATGTGT (residues 187 to 207 of SEQ ID NO:270)

Tm5=58.36

3'primer=TCGTATGGTCGTCAGTGCTCA (residues 233 to 253 of SEQ ID NO:270)

Tm3=58.89

probel=CACCACAAGAGATTGTAC (residues 214 to 231 of SEQ ID NO:270)

probelstart=214

probelStop=231

direction1=Forward

Tm1=68.98

score1=1.98

length=67

WO 01/38351

187/201

PCT/US00/28888

CT628
Nucleotide
Genomic coordinates:
Start: 240139
Stop: 239455 (SEQ ID NO: 272)

Amino Acid
MDSLISKLENIFSIAEQDFFNADSMFMQTMLLPTDAMFTDCESPLYKNKSGGKNIVTDVG
ESVLSSSSDEKMSFKVLSHVLRFPVLLHCNYKQTNTPLWKELYKHGKFALLGDLVLFNS
PFHPNIPAMPFDKSPICDTTGKSIIMSEVMTKELLYKLADKDIGQFFAVLNVTNPITGDS
FLHYFAGGNTMRDGEKICTSADVLRIIAEITIQKTGKMPYELMKK
(SEQ ID NO: 273)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| SYT_HAEIN (P43014) THREONYL-TRNA SYNTHETASE (EC 6.1.1.3) (T | 39 | 0.028 |
| SYT_ECOLI (P00955) THREONYL-TRNA SYNTHETASE (EC 6.1.1.3) (T | 38 | 0.049 |
| CDR1_SCHPO (P07334) MITOSIS INDUCER PROTEIN KINASE CDR1 (EC | 35 | 0.55 |
| Q55626 (Q55626) METHYLENETETRAHYDROFOLATE DEHYDROGENASE | 32 | 3.7 |
| ERG2_SCHPO (P87113) PROBABLE C-8 STEROL ISOMERASE (DELTA-8- | 32 | 4.8 |
| Q85056 (Q85056) SEGMENT 2 | 31 | 8.3 |

Comments:

TaqMan Primer/Probe Sets:
5'start=246
5'stop=267
3'start=288
3'stop=305
5'primer=GCGATTCCCTGTCCTACTTCAT (residues 246 to 267 of SEQ ID NO:272)
Tm5=58.22
3'primer=TCCTTCCACAGGGGCGTA (residues 288 to 305 of SEQ ID NO:272)
Tm3=58.60
probel=CAACTACAAGCAGACGAA (residues 270 to 287 of SEQ ID NO:272)
probelstart=270
probelStop=287
direction1=Forward
Tm1=69.05
score1=1.94
length=60

WO 01/38351

188/201

PCT/US00/28888

CT629
Nucleotide
Genomic coordinates:
Start: 247143
Stop: 244950 (SEQ ID NO: 274)

Amino Acid
MEGGDQRTKLT PATVMGLYQSKTPGEGEGEGGGQFKIPSAIAVKSCCSKNATRRSPPSD
SPYSLRPMKRLKKNNGEVGGKAPPPVTLRLREDYESTPYNFNRNKKRPITIDENOFATL
NPTYATDIIKKQQLPSVSAASVLRKHRANADTQYRKRFSHPNCAKFSTVNLKARDYTPLS
VLRSHVKGPKHLKSSCDVTETNVVKNRFSSIDKWVKLEKPPCYFAVAEADTNIAAGLES
PFHLIRQAAKLGLISDVQDVSSNYETIKQSCIDAKEKASKFLWSNNRTKQPPSSWWPVG
GSKNLSVLDTSPLLNNWRLCKNNGKGWIKTMSIDHMAKNVFKLSPGACESILEKKTLLG
EVT AQCKKWESYRRNIPVPAHVQPEYASQVVMIGPSELYLEVKG VYVMLETGKVIK FMT
DKEMYCEFFVFETVFSHALEGRMKGAVGVRKMCVEGFCVEMDFAGISVIDVLNGDLKCKMD
ENVVQQPNPSTTSSKPAAELMQDHGSLCRMRTLYGVRMLQATGRLPEGLQSKCKKPITD
SISAIAIVGKMRERMLNQLPFVLVEIVNIVTRLSQOGLVNPDIKSDNIVIDGITGQPKMI
DFGLIVPCKKYNFKCWGTDERFFSNHPHTAPEFINSEL CSETAMTFGLAYLLIDMLSIL
IKRTADLSANSIYTNIPFLSIVSKMYDQEKTNRPAYEIA PVIGACFPFKDNI AKLFQSP
KHSLSYKKVK
(SEQ ID NO: 275)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q20443 (Q20443) PRK-2 PROTEIN | 46 | 0.001 |
| Q17735 (Q17735) SIMILAR TO PROTEIN KINASE | 44 | 0.004 |
| KPC1_LYTP1 (Q25378) PROTEIN KINASE C (EC 2.7.1.-) | 43 | 0.007 |
| O01715 (O01715) PROTEIN KINASE C | 43 | 0.010 |
| O01716 (O01716) PROTEIN KINASE C (FRAGMENT) | 43 | 0.010 |
| Q21734 (Q21734) PUTATIVE RIBOSOMAL PROTEIN S6 KINASE II ALP | 42 | 0.017 |

Comments:

EST confirmation of the predicted transcript:
An isolated EST has equence identity to nucleotides 1 to 933 of CT629: this
corresponds to nucleotides 244994 to 245926 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=933
5'stop=954
3'start=1020
3'stop=1040
5'primer=CCCTCTCTTGAAGTGAACAGG (residues 933 to 954 of SEQ ID NO:274)
Tm5=58.83
3'primer=GCTCCAGGGGAAAGCTTAAAA (residues 1020 to 1040 of SEQ ID NO:274)
Tm3=58.49
probel=TGAGCATCGATCACATGG (residues 992 to 1009 of SEQ ID NO:274)
probelstart=992
probelstop=1009
direction1=Forward
Tm1=68.98
score1=1.98
length=108

WO 01/38351

189/201

PCT/US00/28888

CT1030
 Nucleotide
 Genomic coordinates:
 Start: 108549
 Stop: 109164 (SEQ ID NO: 276)

Amino Acid
 MSSGKVITYEIVEGGLLNKKYLLDGGAAICLQSNVCARKRHAGSLHDNLFKMLGFGDPYKQ
 RRGKTNSKNLAIIEDRPQLGSVSVVQHPTPEPERFCSMTFLFAQYNMGNGRKCYPNDKEY
 VESCKKHERVHKSSSTEMKRLRLYYFNKCLHAIKSPAMKKYNKIIFPARIGCAAAGGDWE
 KYHASIRDFSTIIDKEVIIVSQRM
 (SEQ ID NO: 277)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|--|-----------------|------------|
| Q9YTM9 (Q9YTM9) ORF 36 | 32 | 4.3 |
| Q9XAQ7 (Q9XAQ7) NUOD, NADH DEHYDROGENASE SUBUNIT | 31 | 9.7 |

Comments:
 EST confirmation of the predicted transcript:
 An isolated EST has equence identity to nucleotides 1 to 604 of CT1030: this
 corresponds to nucleotides 108546 to 109149 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=284
 5'stop=307
 3'start=387
 3'stop=408
 5'primer=GCTCCATGACATTCTTATTTGCTC (residues 284 to 307 of SEQ ID NO:276)
 Tm5=57.98
 3'primer=TTCTGTGGAAGATTTGTGGACC (residues 387 to 408 of SEQ ID NO:276)
 Tm3=57.62
 probel=TGTTGAGAGCTGCAAGAA (residues 360 to 377 of SEQ ID NO:276)
 probelstart=360
 probelStop=377
 direction1=Forward
 Tm1=69.08
 score1=1.91
 length=125

WO 01/38351

190/201

PCT/US00/28888

CT1031
 Nucleotide
 Genomic coordinates:
 Start: 109260
 Stop: 110088 (SEQ ID NO: 278)

Amino Acid
 MSSNRFSQLRGNEEMVGDYSRWTTVKNRRNRQQQYSHSFRPQQQQQHOKRTSTNSPPAPP
 PPFPIISWGALGSYSMYRLDDQCRNCDETGYYNFHSYDRKRERVRS LNNTPSEGMWRRTS
 RSSPFLNKKKDVDEAPPPQSNQHMYPLNKYSFREYTPSSKLVNWRDPSQEKQDKILQEEE
 ARAPTPTPQEKEPEVETKDDVVIEEETAPEPEPEPAPVDPDPDIPAITATTTTTTVATRHD
 DSSTVFLRNVILSIVFWFLGVYSALFAKCIKRSKKE
 (SEQ ID NO: 279)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q01823 (Q01823) ORF-3 | 48 | 5e-05 |
| PAR1_TRYBB (P08469) PROCYCLIC FORM SPECIFIC POLYPEPTIDE B1- | 48 | 7e-05 |
| Q99356 (Q99356) PROCYCLIN | 48 | 9e-05 |
| P73032 (P73032) HYPOTHETICAL 185.1 KD PROTEIN | 48 | 9e-05 |
| PARC_TRYBB (Q06084) PROCYCLIC FORM SPECIFIC POLYPEPTIDE B-A | 48 | 9e-05 |
| Q27045 (Q27045) SCHIZONT/SPOROZOITE SURFACE PROTEIN (FRAGME | 48 | 9e-05 |

Comments:

EST confirmation of the predicted transcript:
 An isolated EST has sequence identity to nucleotides 137 to 680 of CT1031: this corresponds to nucleotides 109401 to 109944 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=255
 5'stop=277
 3'start=347
 3'stop=364
 5'primer=TTGCGATGAACTGGCTATTACA (residues 255 to 277 of SEQ ID NO:278)
 Tm5=58.44
 3'primer=ATCTACTTGTGCGCCGCC (residues 347 to 364 of SEQ ID NO:278)
 Tm3=58.42
 probel=CTCCAAGTGAAGGCATGT (residues 329 to 346 of SEQ ID NO:278)
 probelstart=329
 probelstop=346
 direction1=Forward
 Tm1=68.97
 score1=1.97
 length=110

WO 01/38351

191/201

PCT/US00/28888

CT1032
 Nucleotide
 Genomic coordinates:
 Start: 119056
 Stop: 121081 (SEQ ID NO: 280)

Amino Acid
 MAGNRTQFVSSLIACISDVEQGMCCGRQAQDALMTRLANLKLGD SLKETDVNLEYLRY
 ASTPLLGE LNYDKQQAATVDINLMAHFSYAALGIESILNSIRRVVANHQRNNGKKPS
 EPISRPHPLGGVEPPLSSELANAI RDKFISMGALORLNSAIVTAALGAIASERELFLREN
 AVNYMYDVEFAERDAATTD TGNVVYLSTKMDEDEDDIKRSEILDKVS KRPAKEGIDWRP
 TPDNSFPYQLIWGDDSVDDTVLIDLITNAIVPNIFMAKFILFICNHLRAVIRSMREILYG
 NISSSSDNYFEDGRKWCFWLNLYNRLEWFMLVVRVFI FLHKKESFSGADNVNVKRLLVV
 VVESFPPVLLDTEWVKTNITSWPVINNSNNNSTLPVTEDTLMRLAIRTSSGARHPIFDEI
 NSLTTAVTNRITFQSAEFCTKILLGRALDEEEAGTKMLVKS VKETGEEKDKNNTFSSFGL
 LLKNTKNEELEINIGDNDDETTDVACWARTSSTSFIRNRTYAFKKIWGLE DASDVVELKR
 ESDAITSFVTDKSSPLLFPYVSDWSCLLHPCKAPAIKSVWLQILKDFSQENIKTINE
 KVQSLSS EICQKSNDRFKNKKIAAEHVRSVKLLNTISNREQEALSTEHC IWLTLWKQ
 VVQNTLN LLENFPV
 (SEQ ID NO: 281)

Top Blast Hits

Sequences producing significant alignments:

| | Score (bits) | E Value |
|---|-----------------|------------|
| 025761 (O25761) HYPOTHETICAL 88.8 KD PROTEIN | 38 | 0.29 |
| 075130 (O75130) KIAA0635 PROTEIN | 34 | 3.3 |
| Q17425 (Q17425) HYPOTHETICAL 62.4 KD PROTEIN B0024.10 IN CH | 34 | 3.3 |
| O33600 (O33600) PURINE NTPASE | 34 | 4.4 |
| Y066_NPVAC (P41467) HYPOTHETICAL 94.0 KD PROTEIN IN POL-LEF | 34 | 4.4 |
| Q53626 (Q53626) ATP-BINDING PROTEIN INVOLVED IN MITHRAMYCIN | 33 | 5.7 |

Comments:

CT1033
Nucleotide
Genomic coordinates:
Start: 121099
Stop: 123634 (SEQ ID NO: 282)

Amino Acid
MFTHLTRA FRKMNNLVNRSFIDVHRVVAELSYPEFEEDVKNPESSIYRTPISL FQNKDIV
TIVGDYILSPKTD SFQVLYPIKKVIEHFPVIFHCTHNNAPLWVHLLDERHHRLLSLLTY
EIVNAKYRGIVVIPYRRPINYQTGKSLMSKLASVKVLDILMRCGSYKFISLMCMINKK
NNTNFLHCCASKWGEVGSKMMLHIAEMFFANPTTSQHLSDASSFPDAAAEDDKGKTPAHL
AIQEDNADALLFLISLYGAPWFQDNN SYMKSAL ELKSNKCVKLSFAADKYEILPNINNN
QLEPDTMCGVCATSV EEDENEGKTTSLSWYQMNC KHYIHCECLMGMC AAAGNVQCPMCRE
DVGDEVLERCPPTIFRWLKLAE RSEHNRVLFEAKKQEFYKQMEAMKPPRVVPPRRTFLT
PARRGERAIRIAREIATNAIAEATAQGDVNSYFPVLIDGSGEYEEEGEEFFNSEEALALA
FGRPFLEDEEEARQIQMRQFAELSRRGVSVNIINNDNPHRHISTVNIVQPVYGEKSPAA
SFIYNMLKNDVFESIRSRDTRVGGERVPMNLSNDKRALFHAASSMLCDFATETNSQIVG
LDFQAVYDPHHISNYIETFGSPLHAYPGA VTFDGAQDYAESIRYDNDIVSFSEMASEL
HITEALDVFEGLLSPLFKKIRTGKSYSNWN DHLRRRNYARDIAEEFVRVCENSLASREH
PPVHVHPPFRDGAIPILIEYIVDFIHCITWSMQVNALHCMRKYIEHENTNVHLLNLRPTD
ERVEVLRVSQLRWSRLFNEQYNTRMSLSTKRLSLMKIFNHD LGVSKFGVYKLLDIEMYC
FTLI
(SEQ ID NO: 283)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| Q9XUM8 (Q9XUM8) W02A11.3 PROTEIN | 39 | 0.17 |
| YEYC_CAEEL (Q21802) HYPOTHETICAL 64.6 KD PROTEIN R07B7.12 I | 38 | 0.22 |
| CAB40955 (CAB40955) HYPOTHETICAL 22.5 KD PROTEIN | 37 | 0.50 |
| YBR2_YEAST (P38239) HYPOTHETICAL 13.2 KD PROTEIN IN ORC2-TI | 37 | 0.50 |
| Q47955 (Q47955) HHDA PRECURSOR | 37 | 0.50 |
| O82239 (O82239) F17A22.9 PROTEIN | 37 | 0.50 |

Comments:

EST confirmation of the predicted transcript:
An isolated EST has sequence identity to nucleotides 1 to 208 of CT1033: this corresponds to nucleotides 123449 to 123656 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=1206
5'stop=1225
3'start=1269
3'stop=1287
5'primer=GGAAGCAATGAAACCTCCCA (residues 1206 to 1225 of SEQ ID NO:282)
Tm5=58.56
3'primer=GATGGCTCGTTCGCCTCTT (residues 1269 to 1287 of SEQ ID NO:282)
Tm3=58.97
probel=TTGTTCTCTCCTCGCAGGA (residues 1232 to 1249 of SEQ ID NO:282)
probelstart=1232
probelstop=1249
direction1=Forward
Tm1=68.99
score1=1.99
length=82

WO 01/38351

193/201

PCT/US00/28888

CT1034

Nucleotide

Genomic coordinates:

Start: 123757

Stop: 126556 (SEQ ID NO: 284)

Amino Acid

MAADLLELAIQETIQSELEEIADTEFLNYLPHKTGICEEAAAANGRPYLPPTLEMRNEVDHF
 WSQDNRKLLKLLGHFCGNLYVEAFIAGSIDAETCVGFLRSQATGLGYPLLKKLALIAREDK
 SNTTNYNLYIDRNSMMKQVFSAEIDKRPSSIQNTSHTKSSPVYLKLIIDRRTECLALDWLD
 ASKRTAKEIGAARKVCFLQNLIVAILIPAYTETFVLDTGNELEQQVLDLDDAYFNAENKDKV
 DEMCVVAILSTLHNLFVRKSLPHHLYNAPFRLPPFGQHPIINIENSSFFNEDTTPILASI
 SIPSSMVIKHHTRKNSRWRCNNLMTAAERSIFLRGVLTVSGDYGWFSVIVGSTIMPSVL
 FYGDRKHLINTVKSNFSAITCSYWNKYMDCRSYGFEIIDTPENNCGRIRAAIDCSNTD
 FHSPVTRVNKKKTSIINAVKNPFFIRHTEPKWYNKNAMCGEVLENVGVTLQHVRSDEY
 MDRFGSLLLGREKKWTCNYLDRIKSLETISNNLKGKIDTMCKILETKYNYKSSSLYYKQI
 TATSDDPKMKIIASINKRRYLCNILEFAIISSEKKDEVEDHTKTGNGGCAFSKYKKKQ
 LEPKQHLIVKVNKYIEAFSLIKMLRND CERNKCRFKEAIRECANELVRELYRASARSYV
 HDLVLKRTNVHLTWQRPYDENANTIMSLIPKCKLHTVLVDKDSRDVKLLNFLTRDGNYN
 PIRHSMLELVYGEYAKDVSTVTCFEWLKWCSSKGVIKYEDFLDRYEKTGEEDKDEREFF
 RLKKCSRDHDTKDIKKIENVLNSDTLYSYSLDKNVQTHASSSTVVKNDDTGKTSVMVGWDYI
 FSIGKGEKTTKKRKLETIDISSDDDDDEEEEEDEGKRMKNMNCSSSIKNKSKNKNGRMC
 CTDILNVVEPSLPNTLSFNCVKMSMDVLNLL
 (SEQ ID NO: 285)

Top Blast Hits

| Sequences producing significant alignments: | Score (bits) | E Value |
|--|-----------------|------------|
| 062235 (062235) F36F2.3 PROTEIN | 46 | 8e-04 |
| AAF04442 (AAF04442) PUTATIVE TRANSLATION INITIATION FACTOR | 43 | 0.012 |
| 060313 (060313) KIAA0567 PROTEIN (FRAGMENT) | 41 | 0.037 |
| Q90491 (Q90491) DNA BINDING PROTEIN E12 | 41 | 0.048 |
| BAA82145 (BAA82145) MYOSIN HEAVY CHAIN 2B | 40 | 0.063 |
| 097303 (097303) PFC1060C PROTEIN | 40 | 0.063 |

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 647 of CT1034: this
 corresponds to nucleotides 125649 to 126295 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=1391

5'stop=1410

3'start=1468

3'stop=1489

5'primer=AAAATGTTGGCGTGACCCTC (residues 1391 to 1410 of SEQ ID NO:284)

Tm5=58.49

3'primer=ACGTCCATTTCTTTTCTCGTCC (residues 1468 to 1489 of SEQ ID NO:284)

Tm3=58.24

probel=AACAACACGTCCGTGTTA (residues 1412 to 1429 of SEQ ID NO:284)

probelstart=1412

probelstop=1429

direction1=Reverse

Tm1=69.00

score1=1.99

length=99

SBV. ORF022.txt
Nucleotide
Genomic Coordinates:
Start: 265664
Stop: 266297 (SEQ ID NO:286)

Comments: EST confirmation of transcript not predicted computationally.

TaqMan Primer/Probe Sets:
5'start=230
5'stop=252
3'start=321
3'stop=343
5'primer=CAAATCAAAGGCCATATGAAACC (residues 230 to 252 of SEQ ID NO:286)
Tm5=58.41
3'primer=GGTGGTCTTTATTTTCATCACCG (residues 321 to 343 of SEQ ID NO:286)
Tm3=58.41
primerScore=0.86
allele1=
probel=GAGGATAGCCAGTTCTAT (residues 265 to 282 of SEQ ID NO:286)
probelstart=265
probelstop=282
direction1=Reverse
Tm1=68.99
score1=1.99
length=114

C003
Nucleotide
Genomic Coordinates:
Start: 126709
Stop: 127126 (SEQ ID NO: 287)

Comments: EST confirmation of transcript not predicted computationally.

TaqMan Primer/Probe Sets:
5'start=175
5'stop=194
3'start=265
3'stop=283
5'primer=CATGGTGTGCCTAAAGCAGG (residues 175 to 194 of SEQ ID NO:287)
Tm5=57.84
3'primer=TGGCACAATCATGGGTGA (residues 265 to 283 of SEQ ID NO:287)
Tm3=57.81
primerScore=0.91
allele1=
probel=AATGGAGGCATCGAAGAA (residues 212 to 229 of SEQ ID NO:287)
probelstart=212
probelStop=229
direction1=Reverse
Tm1=69.02
score1=1.97
length=109

C010
Nucleotide
Genomic Coordinates:
Start: 184094
Stop: 184931 (SEQ ID NO: 288)

Comments: EST confirmation of transcript not predicted computationally.

TaqMan Primer/Probe Sets:
5'start=547
5'stop=570
3'start=655
3'stop=678
5'primer=AAGATGATGAAACCTGGTTCAAAA (residues 547 to 570 of SEQ ID NO:288)
Tm5=57.75
3'primer=CGTCATGGAAGAAAAGGAATTGTA (residues 655 to 678 of SEQ ID NO:288)
Tm3=58.24
primerScore=0.67
allele1=
probel=CTCCTTGCAAGTGCGATA (residues 598 to 615 of SEQ ID NO:288)
probelstart=598
probelstop=615
direction1=Forward
Tm1=69.01
score1=1.98
length=132

C020
Nucleotide
Genomic Coordinates:
Start: 60122
Stop: 60341 (SEQ ID NO: 289)

Comments: EST confirmation of transcript not predicted computationally.

TaqMan Primer/Probe Sets:
5'start=2
5'stop=26
3'start=118
3'stop=140
5'primer=CAACCTCCACATCTCCTTCTAATTC (residues 2 to 26 of SEQ ID NO: 289)
Tm5=58.27
3'primer=AACGCTTTGTTTAACGTGCTTTT (residues 118 to 140 of SEQ ID NO: 289)
Tm3=58.18
primerScore=0.61
allele1=
probel=ATTCAAGAGGTCACAAACA (residues 80 to 98 of SEQ ID NO: 289)
probelstart=80
probelstop=98
direction1=Forward
Tm1=68.75
score1=1.50
length=139

WO 01/38351

198/201

PCT/US00/28888

C028
Nucleotide
Genomic Coordinates:
Start: 277009
Stop: 277324 (SEQ ID NO: 290)

Comments: EST confirmation of transcript not predicted computationally.

TaqMan Primer/Probe Sets:
5'start=293
5'stop=315
3'start=384
3'stop=406
5'primer=GGTGCAGTTTCATCCTTACCATT (residues 293 to 315 of SEQ ID NO: 290)
Tm5=58.07
3'primer=TCGTCATCATAGTCGTCGTCAAC (residues 384 to 406 of SEQ ID NO: 290)
Tm3=58.72
primerScore=0.86
allele1=
probel=TTCCACAACCACCACTAC (residues 320 to 337 of SEQ ID NO: 290)
probelstart=320
probelstop=337
direction1=Forward
Tm1=68.94
score1=1.94
length=114

WO 01/38351

199/201

PCT/US00/28888

C036
Nucleotide
Genomic Coordinates:
Start: 115067 (SEQ ID NO: 291)
Stop: 115380

Comments: EST confirmation of transcript not predicted computationally.

TaqMan Primer/Probe Sets:
5'start=99
5'stop=117
3'start=158
3'stop=177
5'primer=CCACTGCTGGACGCATCTC (residues 99 to 117 of SEQ ID NO:291)
Tm5=58.94
3'primer=CGACGGACAGTGGAGCTCTT (residues 158 to 177 of SEQ ID NO: 291)
Tm3=59.04
primerScore=1.25
allele1=
probel=CCATCAATGAAGAAGCGT (residues 131 to 148 of SEQ ID NO: 291)
probelstart=131
probelstop=148
direction1=Forward
Tm1=68.97
score1=1.97
length=79

WO 01/38351

200/201

PCT/US00/28888

C037
Nucleotide
Genomic Coordinates:
Start: 186828
Stop: 187047 (SEQ ID NO: 292)

Comments: EST confirmation of transcript not predicted computationally.

TaqMan Primer/Probe Sets:
5'start=63
5'stop=82
3'start=133
3'stop=157
5'primer=TTGACATCAGACCGACCCAG (residues 63 to 82 of SEQ ID NO: 292)
Tm5=58.08
3'primer=GGAATTTTACTCTTCCTCCATCTCA (residues 133 to 157 of SEQ ID NO: 292)
Tm3=57.92
primerScore=0.62
allele1=
probel=CACCCTCTAAACTCGAGC (residues 88 to 105 of SEQ ID NO: 292)
probelstart=88
probelstop=105
direction1=Forward
Tm1=69.49
score1=1.50
length=95

WO 01/38351

201/201

PCT/US00/28888

C019
Nucleotide
Genomic Coordinates:
Start:141541
Stop: 141647 (SEQ ID NO: 293)

Comments: EST confirmation of transcript not predicted computationally.